

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGGCAGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCTGTCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTTC
CTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGA¹CTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTTTTCCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAA
CGTGAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCVPGGGPPRVAPNPVTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCCCTCGCCTCCCCGCAGAGTCCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCACCGCGGGGACTATGGTGAAATTCCCGGCGCTCACGCACTACTGGCCCCCTGATC
CGGTTCTTGGTGCCCCCTGGGCATACCAACATAGCCATCGACTTCGGGGAGCAGGCCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGAAGTTCAAAAATGTGGGCCTGGTGT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTACAGACATCATCCC
TGACAGAAGTGGCCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCAGCAGCAACACAGTCACGGCAGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGCG
GATCATCGTCTCATCGCCAGCCTCGTGGTCTTACCCTACCTGGGGGTGCACGGTGCAGCCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTGTTTTGTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCAGCTGCAGCCCGGAGTCCCCGTTACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACCTGACATGGTTGAATGATAGCATTTCCTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTGAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTCACAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTCGCATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACCTGAAGCTTTAAGTCTCATCCAGCATTTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACACCATAACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACATATGTGGTCCCTCCCTGAAATAACGCATTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTTGAAAGCAGGTTTTTTCTTTAAAAAAATATATAGACACGGTTCACT
AAATTGATTTAGTCAGAAATTCCTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSTDFKNVGLVFNNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYIINKLHHVDESV
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLC LGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVVPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVBHATLGVBGSLL
AGFVGESTMVAIAACYVYRKQKKKMENESATEGEDSAMTDMPTTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNAAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

[illegible]

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAACAGTNACGGGCAGCCACATCAAGAAGTTTACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTTCGTGATGTTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTTCTTTGCGGATTTTTCT
CCTTTTTTCCAGTTCAGTCAACAGNGAGGGCGCATCTCACCGGGNNGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCTTNATNGCCAGCCTTGT
GGTCCTACCTTACCTGGGGGTGCACGGTGCGACCTGGGCGTGGGTTCCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.15.438101>; this version posted April 15, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCGGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTTACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCTTCTTCA
TCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTATGTACTACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCCTGCCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCCTCCTCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACTCCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT
CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCTTGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW
NQRWLKGAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVITYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICAS
WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAGGCCATGGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGA CTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTTGGCTGGG
TGTATGGGGGAATAACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGTATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAATTGCAGGA
GCTGTCACGGGAAGTCTTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGCTTTTTTCTTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAETIYHNRFDVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTTCGTTTCATG
GCTGGCGCCGAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted September 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GAGCCGCCGCCGCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCGGCCCGCCACCCACGTCT
GCGTTGCTGCCCCGCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCTAGGGAACCT
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTGCGGGG
GCTTCGCGTGTTC AAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTAT
GTTT CAGTTTTCTGTATCTTGCCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAACTGCT
GTGGGTTCCGAAGTGTTAACC CAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCGAATCCTAGTGCATTCTTTGA TGAGAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT
CTATGTTTTCTCTACATGTTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTGCGGC ACTGTCCACTGTGGCCTT
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACC ACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCTCTAC
TGGAAAAAGAGTGGAATTTATTAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCCAATTTTTTTTTTGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGTTTTAATTTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCC CACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATCTGT TAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTC CGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

[illegible]

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

[illegible]

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCTTGGTGTG
CTCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCCTCCACCTGATTGTGTCCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCTTGGGATCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCTGCAGAGGTGAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCACTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGGTGCTTCCAGCTCT
GCTGCACCTCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCTTGGGCCTCTGCCTCCGCCGCCAACCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCTCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGCTCCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAACGAAAAATAATCTGGCCCAAAATGTCAGTTGTACTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATAACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTCGCCAAACACCG
ACTCTGTGCTTGCCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSAAPPPTLNPAQKSAAPGTAPEEAPGPLPPPPTPSSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAAGAAAGTCTCTCCTCTTCCATTTTTTGACCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCCCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLLTVDTYGRPILEVPESVTGPWKGDVNLPCITYDPLQGYTQVLVKWLVRGS
DPVTIFLRDSSGDHIQQAKYQGRLLHVSHKVPBGDVSLLSTLEMDDRSYHTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRLSLQCQARGSPPISYIWYKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTMT
YPLKATSTVKQSWDWTDDMDGYLGETSAGPGKSLPVFAILLIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

protein = seqdb2

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTGCGGA
GGCGCCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGGATGTCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCTGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGGAGTAACGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCTCTTGTTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACACACCCAGCATGATCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCTGAATTACAATGGACTTGACTCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCTTATAACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAATTTCTATGCATTTCTGCAAACTTATTGGATT
ATTAGTTATTGAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTGATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
CAGATGCCTCTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
TATCAACGTCTTAGAAAGAAATCTTCTAGAGAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATAACATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

[illegible]

```
><subunit 1 of 1, 373 aa, 1 stop
```

><MW: 41281, pI: 8.33, NX(S/T): 3

Signal sequence:

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTCGTTCTTTGCTCTCTCGCGCCAGTCTCTCTCCCTGGTTCTCTCAGCCGCTGTGCGAGGAGAGCACC CGGA
 GACGCGGGCTGCAGTCGCGCGGCTTCTCCCCGCTGGGCGGCTCGCCGCTGGGCAGGTGCTGAGCGCCCCTAG
 AGCCTCCCTTGCCGCTCTCTCTCTGCCCCGCGCAGCAGTGACATGGGGTGTGGAGGTAGATGGGCTCCCCG
 GCCCGGAGGGCGCGGTGGATGCGGCGCTGGGCAGAGAAGCAGCCGCGGATTCCAGCTGCCCCGCGCCCCGGCG
 CCCCTGCGAGTCCCCGGTTTCAAGCATGGGGACCTCTCCGAGCAGCAGCACCCGCCCTCGCCTCTGCGAGCCGATC
 GCGCGCCGAGCCACAGCCACGATGATCGCGGGCTCCCTTCTCTGCTTGGATTCTTAGCACCACCACAGCTCAG
 CCAGAACAGAAGGCCCTCGAATCTCATTGGCACATACCGCCATGTTGACCGTGCCACCCGCCAGGTGCTAACCTGT
 GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCGAGCAGTTGCCCT
 GTGGGGACCTTTACCAGGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCATGGCCAATG
 ATTGAGAAATTACCTTGTGTGCTTGTGACTGACCGAGAATGCACCTTGCCACCTGGCATGTTCCAGTCTAACGCT
 ACCTGTGCCCCCATACCGTGTGTCTGTGGGTTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
 TGTAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT
 CTGAGTCAGAACCTGGTGGTGTATCAAGCCGGGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCTTTC
 TCCAGCTCCACCTCACCTTCCCCCTGGCACAGCCATCTTTCACGCCCCGAGCACATGGAAACCCATGAAGTCCCT
 TCCTCCACTTATGTTCCCAAAGGCATGAACCTCAACAGAATCCAACCTCTTCTGCTCTGTTAGACCAAAGGTACTG
 AGTAGATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCTC
 CCAAACCTTTCAGGTAGTCAACCACAGCAAGGCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCATGGAG
 GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG
 CATTGACATCAATGAGCATTGTGCTTGGATGATTGTGCTTTTCTGCTGCTGGTGTGCTGTGGTGTGTTGGTG
 TGCAGTATCCGAAAAGCTCGAGGACTCTGAAAAGGGGGCCCCGGCAGGATCCAGTGCCATTGTGGAAAAGGCA
 GGGCTGAAGAAATCCATGACTCCAACCCAGAACCAGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
 ATCTGTAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG
 AGGGAGGTTGTGCTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCTACGAGCTCTGCGAGCTGAGGACC
 ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCCGAGAAACGATGTTGTGGAG
 AAGATTTCGTGGGCTGATGGAAGACACCACCCAGCTGGAACCTGACAACTAGCTCTCCCGATGAGCCCCAGCCCCG
 CTTAGCCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCGCTCTCTGACGGTGGAGCCTTCCCCA
 CAGGACAAGAACAAGGGCTTCTTCGTGGATGAGTCGGAGCCCCCTCTCGCTGTGACTCTACATCCAGCGGCTCC
 TCCGCGCTGAGCAGGAACGGTTCTTATTACCAAAGAAAAGAGGACACAGTGTGCGGCAGGTACGCCCTGGAC
 CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTCTAAATCCTGAGGAGCTGCGGGTGTGTAAGAG
 ATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTGCAAAATTATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
 CTCCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAATTACTCA
 ATTTAGTGGCAGGGTGGTFTTTTAAATTTTCTTCTGTTTCTGATTTTGTGTTTGGGGTGTGTGTGTGTGTTGT
 GT
 TCTCTCTCTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAAGTGTGTGAA
 ATACCCACCACTAAAGTTTTTAAAGTTCCATATTTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTCTG
 TGCATTTTAAATTTACTTTAACTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC
 TTCTTAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTATGTCTCTTAACATTACACCTACTTTTT
 AAAAACAATATTATTACTATTTTATTATTGTTTGTCTTTTATAAATTTTCTTAAAGATTAAGAAAATTTAAGA
 CCCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTTCATATTTCATGG
 CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTTACCCTGGACACCGTGTAGAATGCTTGATTACTTGTAC
 TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT
 GACAACCTGGGCCACCAAGAACTTGAACCTTCACTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACCTT
 GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCAGCTTTGCTTTAAAGATGTCTTG
 TTTTTTATATACACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCTGCTGGGATTCTTCCACCAATT
 ACTTTAATTAATAATGGCTGCAACTGTAAGAACCCTTGTCTGATATATTTGCAACTATGCTCCATTATCAAAATG
 TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTGTGG
 GTAGTGGTGAAGGACCGATATCAGAAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTA
 AAAAAAAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

><subunit 1 of 1, 655 aa, 1 stop

><MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIQTYRHVDRATG
QVLTCDKCPAGTYVSEHCTNTSLRVCS SCPVGTFRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAIFFRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPLNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGRHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTQNRKWIYYCNGHGIDILKLVA AQVGSQWKDIY
QFLCNASEREVA AFSNGYTADHERAYAALQHW TIRGPEASLAQLISALRQHRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPSPSPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL
RCDSTSSGSSALS RNGSFITKEKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAAATGCCGTGCTCCAGGTGTTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCACTGCACAGCCT
GTGGTCAAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCAGGGCTACCACCTGTGCGGGGGCTCTGTCTACAC
GCCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCCTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCGTGCTCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACCTCCTTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAAC**TGAA**
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAACTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGGAGGTGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCCGCCACCACACCCAACTAATTTTTTGTATTTTATAGTAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGTGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAGAGCAGCAACTTGCAAGGGCGGCCTTTCCCACTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCAGAAAGTGCAGAACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCAAAACCCACCTTTCTACTTCCAAGACTTAT
TTTACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGACGTATTATTTGTCTTTGATTCCAAATAATATGTTTTCTTCCCT
CATTTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLOVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIHLLPDDKVLTALHH
SVYVREGCASGHVVTLOCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIIITAAHCVYDLYLPKSWTIQVGLVSLDNPAPSHLVEKIVYHISKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHR
DVGGGIISPSMLCAGYLTGGVDSCQGDSSGGLVCQERRLWKLVGATSFQIGCAEVNKPQVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

protein sequence

Year	1992-93	93-94	94-95	95-96	96-97	97-98	98-99	99-00	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30	30-31	31-32	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40	40-41	41-42	42-43	43-44	44-45	45-46	46-47	47-48	48-49	49-50	50-51	51-52	52-53	53-54	54-55	55-56	56-57	57-58	58-59	59-60	60-61	61-62	62-63	63-64	64-65	65-66	66-67	67-68	68-69	69-70	70-71	71-72	72-73	73-74	74-75	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	96-97	97-98	98-99	99-00	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30	30-31	31-32	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40	40-41	41-42	42-43	43-44	44-45	45-46	46-47	47-48	48-49	49-50	50-51	51-52	52-53	53-54	54-55	55-56	56-57	57-58	58-59	59-60	60-61	61-62	62-63	63-64	64-65	65-66	66-67	67-68	68-69	69-70	70-71	71-72	72-73	73-74	74-75	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	96-97	97-98	98-99	99-00	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30	30-31	31-32	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40	40-41	41-42	42-43	43-44	44-45	45-46	46-47	47-48	48-49	49-50	50-51	51-52	52-53	53-54	54-55	55-56	56-57	57-58	58-59	59-60	60-61	61-62	62-63	63-64	64-65	65-66	66-67	67-68	68-69	69-70	70-71	71-72	72-73	73-74	74-75	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	96-97	97-98	98-99	99-00	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30	30-31	31-32	32-33	33-34	34-35	35-36	36-37	37-38	38-39	39-40	40-41	41-42	42-43	43-44	44-45	45-46	46-47	47-48	48-49	49-50	50-51	51-52	52-53	53-54	54-55	55-56	56-57	57-58	58-59	59-60	60-61	61-62	62-
------	---------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----

CCACGCGTCCGTCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTATTATGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCACGCACGCACACACGCGGGGAACTTTTTTAAAAATGAAAGGCTAGAGAGTA
GCTCAGCGCGCGCGCGCGCGTGC CGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA
CGCCCGGCCCCGGCTCGGCGCCCCGCTGGGATGGTGACGCTCGCCGCCGGGCCCCGAGAGCTGCTGCACTGAAG
GCCGGCAGCAGTGGCAGCGCGCCCGCTGCCCCTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGCTCT
GCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTGAGTGCCCTCTGT
TCGGAGTGGGACCTCTGGATCCCAGTGAAGAGCTTCAGTCCAAGAATCATCCAGAAGTGCTGAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTATTGCCAGAGTTTCCAGGAA
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTGAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAATGA
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGATCATGTGGATCACATACAACACACAAACCTCGCTGCAAGAATGTGTTTCCACCACCCTCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
AGAGTTTTCAGAGGCAAGGAAAAGACTCTGGAAGAAAGTTAAGCAGCGATTATAGAGATTGCTAATCAGGTTGACAA
GTTTTACAGACCACTGAACATTCCGATCGTGTTGGTAGGCGTGGAAAGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCCATTACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTTGTGAGTGGGGTTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
CACGGCAGACAGTCTGGGGGAATTGTGATGGACCATTGACACAATCCCTTGGTGCAGCCGTGACCTGGCACA
TGAGCTGGGCCACAATTTCCGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGAGGCTGCATCATGAACGCTTCCACGGGTACCCATTTCCTGCTGGTGTTTCCAGGTTGACAGGAGGACTT
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCTGTTTAACTGGCGAAGTCAGGAGATCTTTGGGGGGC
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCCGTG
CTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC
TCACTGCCACGCCAATGTGTACCTGACAGTATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCAGCAGCAGTGTGTGTCAGCTCTGGGGACCAAGTGTCTAAACCTGCCCTGGGATCTGCTTTGA
GAGAGTCAATTCTGCAGGTGATCTTTATGGCACTGTGGCAAAGTCTCGAAGAGTTCCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAAGAAATCCAGTGTCAAGGAGGTGCCAGCCGCGAGTCATTGGTACCAATGCCGTTTC
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGATTTCTGTGCCGGGGGACCCACGTGTACTTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGAGATGGAAGAAATCTGCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTTACGAGTGTGCAATGCAGTGCCACGCGAGAGGGGTGTGCAACAACAGGAAGAA
CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCCC
CATCCGGCAAGCAGGACCAAGGCAGGAAGCTGCAGAGTCCAACGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCGCTCCCATTGACATGGAGACCGTGACCAGTG
CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCAGTCTGCAGAAA
CAGTGC AAGGAAGGGCAGCGACTTCTGTTGAGCTTCTGCTAAAACATGGACATGCTTCAGTGCTGCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTTC
ACTGAGCCTCCACAGCAGTGTGGGGGAGAACAGGGTTGGGCCAGGTGTCCTTCCCTTTCCCAAGTACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCACTTAATGCTCTGATATGGCTTTAGCAATTATATGAAAT
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAAACATGAATGGCAA
TGAAACAAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC
AGTACTCAGGTTTGGGGTTTGCAGAAAGCCAGGGGAACCCACAGAGTACCAACCCCTTCATTTAACAAGTAAGAA
TGTTAAAAAGTGAAAAAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVVSASVRSGDLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPFPPSQTWARRHKRETLCATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWVNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFNMHDTLDRGCSCQMAVEK
GGCIMNASTGYPPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFQQKCGNRFVEEGEE
CDCGEPEECMNRCNATTCTLKPDVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVGVCYNGICQTHEQQCVTLWGPAGKAPAGICFERVNSAGDPYGN
CGKVSKESSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLOQGGRIICRGTHVYLG
DDMPDPGLVLGATKCADGKICLNRCQCNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAESNRERGQGEFVGSQEHASTASLTLL

Signal peptide:

amino acids 1-28

10401: seqdb2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

[The following text is extremely faint and largely illegible due to low contrast and blurring. It appears to be a continuation of a list or index.]

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCGCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTAAAGAAGTTAAT
GAAACCATAACCTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTCGAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTGTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCGTGACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCACTGGCT
CACGCCTGTAATCCCGGCACCTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCACTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVLLAL
LHLYH

Signal peptide:

amino acids 15-27

protein = seqdb2

FIGURE 34

CCCCGGCGAGAGCGCGCCCGAGCCCCGCGCGATGCCCCGCGCGCCAGGACGCCTCCTCCCGCTGCTGGCCCCGGC
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGGGCCATGGCGGCGGGCGGCTGGGGCGCCCGGGGCCAGG
AGGCGGCGGCGGCGGCGGCGGACGGGCCCCCGCGGACAGCGGCGAGGACGGACAGGACCCGCACAGCAAGCACC
TGTACACGGCCGACATGTTACGCGACGGGATCCAGAGCGCCGCGCACTTCGTGATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGCTGCAGCCGACTTGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGGTGCAGGATACCCACCTTAA
AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCTCGGGACTTCAGACACTGGGAAACTGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGAACCGCCAGTGCCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCAGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCCTTGAACATTCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGAAACCAGGTTCTGGCTATCCCACTC
TTCTCTGTTCCGAGATGGGAAAAGGTGGATCAGTACAAGGGAAGCGGGATTTGGAGTCACTGAGGGAGTACG
TGGAGTCGCAGCTGCAGCGCACAGAGACTGAGCGACGGAGACCGTACGCCCTCAGAGGCCCGGCTGCTGGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGCAGAAGGAA
TAACCTTCATCAAGTTTATGCTCCATGGTGTGGTCACTGTAAGACTCTGGCTCTACTTGGGAGGAACCTCTCTA
AAAAGGAATTCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA
AGTATTCGGTACGAGGCTACCCACGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACCTTAGGAACACAGTTGGAGGTCAC
CTCTCCTGCCAGCTCCCGCACCCCTGCTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCAGTGGTGGCT
GTTCAAGAACGAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTCCAGCCAAACACACTCTACAG
ATTCTTTTATTAAGTTAAGTTTCTTAAGTAAATGTGTAACCTCATGGTCACTGTGTAAACATTTTCACTGGCGATA
TATCCCTTTTGACCTTCTCTTGATGAAATTTACATGGTTTCTTTGAGACTAAAATAGCGTTGAGGGAAATGAAA
TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC
CCACGAGTTCTGGAAGGTGGCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTCACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCCCTGAGTAAGAGAATGCTGATGCCATACTTTATGTGTGATACCTTGTCAAATCAGTTACTGTTCAAGGGAT
CCTTCTGTTTCTCACGGGGTGAACATGTCTTTAGTTCTCATGTTAACACGAAGCCAGAGCCACATGAACTGT
TGGATGTCTTCTTAGAAAGGGTAGGCATGGAATTTCCACGAGGCTCATTCTCAGTATCTCATTAACCTCATGGA
AAGATTCCAGTTGTATTTGTCACTTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCTTTTAGAGG
CTTGCTATACTTGGTCTGCTTCAAGGAGGTCGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC
CAAAGACAGATGTCACTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA
ACTCATGCTGCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCATAATAACCACTTTGCATCCAACACTCTTCACCCACCT
CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCCTTAAAGGAAATCTTTATTAATCAG
TATGGTTCACAGATAATTCTTTTTTAAAAAAACCAACCTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTACA
CACAACCTCAGCTTTGCATCACAGTCTTGTATTCCAAGAAAATCAAAGTGTACAATTTGTTTGTATTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTTAA

Parameter	Value	Standard Error	t-Statistic	p-Value	95% Confidence Interval
Intercept	1.0000	0.0000	1.0000	0.3173	0.9999 - 1.0001
Age	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age squared	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age cubed	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age quartic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age quintic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age sextic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age septic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age octic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age nonic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age decic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age undecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age duodecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age tredecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age quattuordecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age quindecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age sexdecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age septendecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age octodecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age novendecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age vigintic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta squared	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta cubed	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta quartic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta quintic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta sextic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta septic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta octic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta nonic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta decic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta undecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta duodecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta tredecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta quattuordecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta quindecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta sexdecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta septendecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta octodecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta novendecic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta vigintic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta unguicquaginta	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta unguicquaginta squared	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta unguicquaginta cubed	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta unguicquaginta quartic	0.0000	0.0000	0.0000	0.9999	-0.0000 - 0.0000
Age unguicquaginta unguicquaginta quintic	0.0000	0.0000	0.00		

```
><subunit 1 of 1, 432 aa, 1 stop
```

MPARPGRLLPLLARPAALTALLLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHTGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFEHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDLSLHRFVLSQAKDEL

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACCAAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTA
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAAATCCCTAACCCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTA^{ACTTCCCTCATCCT}
ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

[illegible]

><subunit 1 of 1, 277 aa, 1 stop

MNGFASLLRRNQFILLVLFLQLIQSLGLDIDSRPTAEVCATHTISP GPKGDDGEKGDPEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGGRY
RKFGVQLDISIARLKTSMKFVKNVIA GIRETEEFYIIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

amino acids 1-25

FIGURE 38

GGTTCATATCGATTTCGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCCTGTGGAGTGGTGGGGGCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTCACCTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTCTGTGGCTCTG
GTTTCGGGCTGGCCAAGGCCGCTGCGCACTGCCTTTGTGCCACCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCGAGCTGCGGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCCTCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCCAC
TCTACCACATGTCCGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGCCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCCTGGCTTTACAAGC
ATATCTTCCCCCTTCTCCTTGATTGCTATGATGTCAACCAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCTCACAACCTGCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGGG
CCGTTGCAGGTGTACTGGGCTGTCAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTGACCTAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCAACTTGTTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLLPL
LLLKLHLWPQLRWLPADLAFVAVRALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRRFSSYSEAERESNRAARAFLRALGWDWGPDGGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGVPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVR LAVGSGLRPD TWERFVRRFGPLQVLETYGLTEGNVATINYTGORGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRD PQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQ GK
LLKD VFRPGDVFFNTG DLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALD FLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLP PYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLS DPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATAACACTCGCTCTC
GCTTGTCATCTCCCTCCCGGGGAGCCGGCGCGCTCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCCGAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTTAAACACTTCTTTTCTTCTCTTCTCGTTTTTGATTGCACCGTTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTTGGCTTGCCATCGT
CCATCTGGCTTATAAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACCAATGCCTTCTTGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTGAGCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTGAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTGAGGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTGAGCCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTGCGGGGCTTCCCACTGTGAGGCCCTGCA
ACAATACTGTCTCAACGTCATGAAGGGCTGCTTGGAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTGAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTTGGTCAGATGAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCGTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTCTCATTTAAAT

><subunit 1 of 1, 555 aa, 1 stop

MPSWIGAVILPLLGLLLSLPAGADV KARSCEV RQAYGAKGFS LADIPYQEIAGEHLRICPQ
EYTCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLEN A E KSLNDMF
VRTYGM LYM QNSEVFQDLFTELKRYT TGNVNLE EMLNDFWARLLERM FQLINPQYHFSE DY
LECVSKYTDQLKPFQDVP RKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGP FNIES
VMDPIDVKISEAIMNMQENSMQVS AKVFQCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGH SKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMMDDVCPTEFEFVTT EAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:
amino acids 1-23

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACCTACGCAAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGGC
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCCTCACTTCGCCGCTGGTGAGTGTGGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCCGCCCTCCGTACCAGCACTCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCTGTCCACTGTGATTCTCAATTCTTGGCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTCTCTTTTTTTTTTCTTGGTGGAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTCTTGCCAGGTGGAGAGTCTTCACCGTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGGCGGGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA
GGGGCTGTGCGGGGCATCCGCCTCCGCCTTCTCCACAGGCCTGTGTCTGTCTGGAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTGCGGGCCAGGCCTTAGAAGAGGAGG
AAGAAGGGGCCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCAGCCCCATCTCA
TTTTCATCTAGCGGATGATCAGGGATTTAGAGATGTGGGTACCACGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTTAC
CTCTGGACAATGCCACCCTACCTCAGAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGAAAATGGCACT
TGGGTTTTAACAGAAAAGAATGCATGCCACCAGAAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCATA
ACCCACAAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCACTCACCAGTCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCAATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC
AAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTCAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCAGATTCAACAAAAC
TGCAGTGCCGGTCAGGTATCCCCCAAAGACCCAGAAGTAACCTAGGCTCAATGGAGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAGCCAAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAGCAAAAAAAA
GAAGAAGAAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCCACTGAAAACTGTCTGCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTGAGTTCCACTGCTGTGCTTCACTCAACTGACCAAACACTGCTTTGAATTATAGGAGGAGAAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAACTACCTTTGATAAATTAC

[illegible]

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVG YHGSEIKTPTLDKLA AEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTH MVGKWHLGFN RKEC
MPTRRGFD TFFG SLLGSGDY YTHYKCDSPG MCGYDLYENDNAAWDYDNGIYSTQMYTQ RVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAML SCLDEAINNVTLA
LKTYGFYNN SIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVG FVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTGAGCCT
GGGGTCTGTCACTATGGAACATAAAGTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACCTTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAAAAATGTTACCCCAAGACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGAAGTGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTACAAGAAAGACATTGGCCGATTGAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAATTGAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTTTGAAGCAGAACGTGGCAAGGGCAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTGAGTTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTGAATTAAGTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATG
TAAGATGCCTTTCTTGATATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPGCKFGECVGPKNKRCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSKYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGCTGTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAATC
TTTAAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAATAATAAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

CGACGTGGGCACCGCCATCAGCTGTTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCCTGTCTACCTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCAGTGGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCACTGCTGCCCAGTCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCTGGGTTCTCTGCAGCGTGAAGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAAGTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTTCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
CTGCCCAGTCTTCAATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAGTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCAGTGGTG
CATGAGGTGAGGGGCACATGGTTTCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTACC CGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACCAGAGCCCAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCTGTCTCCCCACCCCTGTGATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCCCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCTCTCTCTTTACTCTTTTCAATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTCTTTGGGGGGCAGCAGTTTTCTTTTAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHVWQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG
WVLGRARPGAGISSLQTPVPTLLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGGCCCCCATTGCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCCC
TCTTCCCGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCCGGCAGCGCTGCCGTGAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCCTGACTGGCAACGTGACCACGTGCGCAGATGGCCAACGA
GCAGGGCTTGTGTTGATGTGCACAGCGTCCTGCGGGTGGTGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCAGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCAACATCACAGGG
CAGCCTATGACATTTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTCTGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTTACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTTT
CCCTCCTTCTCCTCCCTGCCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGTCTATTCTGGGGCTGGGGCAGTCTTTTCTGGC
TTGCCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEKSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACCTGTCTGAAGCCCTTGTCCGTAAAGCCTTGAACCTACGTCTTAAATCTATGAAGTCG
AGGGACCTTTCGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCTCCTGC
TCTGGGCTATGAAAAGAAGGAAAAGGTAAAATCCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTGCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTAAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTTTATATATTTTTT
CTGACTCCTATTTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTCACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTTATGTGTGAGTTTCTTTGTAATAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCTGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194

><subunit 1 of 1, 211 aa, 1 stop

><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTLFVTSLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGS
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLAIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

[illegible]

ATAGAGAGACTCCATCTCAAAAAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
 GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA
 AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
 TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCCTCAACGTCC
 TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
 AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
 TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
 GAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
 TTTACAACCTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCCGCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACTGGC
TCTGAGGAGATGTTCCCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCC

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA

FIGURE 58

TGCGGCGACCGTTCGTACACCATGGGCCTCCACCTCCGCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGACGTCACCCCCAGTGGTGTGCTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAAGTGTGCTGCTGCTGCTCATCATTTGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCTGGACCCAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTGC
TGGTTGCCCCACAGTATGGGCAACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGAGGCCTGGAAGGACAAGT
ATATCCGGGCCTTCTGTGCTCACTGGGTGCGCCCTGGGGGGGCGTGGCCAAGACCTGCGCGTCTTGGCTTCAGGAG
ACAACAACCGGATCCCAGTTCATCGGGCCCCGTAAGATCCGGGAGCAGCAGCGGTGAGCTGTCTCCACAGCTGGC
TGCTGCCCTACAACATACATGGTCACCTGAGAAGGTGTTCGTGCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCCTGCTATGGTACTGGCGTCCCCACACCAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAAGTGAAGAGTGGCC
TGCAAGTGGCCAGGCTGGCAGAGCCGCGAGGAGCAACCAAGTGTGTGCTGCAGGAGCTGCCAGGCAGCGAGCACATCG
AGATGCTGGCCAACGCCACCACCTGGCCTATCTGAAACGTGTGCTCCTTGGGCCCCGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTGATGGCCACGCGTTTTGCAAAGTTTGTGA
CTCACCATTCAAGGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGTGTTTTGTTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGGGA
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTGGTCCCAGTCCCTGCCTGGGGCCATG
TGTCCCCCTATTCTGTGGGCTTTTTCATACTTGCTTACTGGGCCCTGGCCCCGAGCCTTCTATGAGGGATGTT
ACTGGGCTGTGGTCTGTACCCAGAGGTCCCAGGGATCGGCTCCTGGCCCCCTGGGTGACCCCTTCCACACACCA
GCCACAGATAGGCCTGCCACTGGTCATGGGTAGCTAGAGCTGTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGAGTTTGTGCGTTCTTCGTGGTTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGTG
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTACCCCTGGGACTGTGTTCCAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA
CACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCACAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCCAGCTTAGTGTGGGACTAGCCAGAACTTGAATGGGACCTGAGAGAGCCAGGGGTCCCC
TGAGGCCCCCCCTAGGGGCTTTCTGTCTGCCCCAGGGTGTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCCTTCATGGCAGTAGGCTTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGTACAGCCTCTAGGT
GGGGTTCCCAAAGACGCCCTTCAAGGCTGGACTGAGCTGCTCTCCCAAGGGTTTCTGTGCACTGGATTTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCGGCAATAAAAGTACTCTGGATGCTGTAAAAA

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSVSTSWLLPYNYTWSPEKVQVQTPTINYTLRDYRKFFQDIGFEDGWL
MQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAAGTGCCA
GCCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCCTCGCTTTTAATGA
CCTCAGCCCCGCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCCGGGTGAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGTCTCATGGCACTTCCTCCTTG
CTCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIQDLLFSALWTFLWVVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 61

FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
CCCGTGGCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCCGGGGTACCTCCGCCTGGTGCCCTGTTTGTGCTGCTG
GCCCTGCTCGTGTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCTAGGGTACAAGGCGGAGGTGATGGTCAGC
CAGGTGTACTCAGGCAGTCTGCGTGTACTCAATCGCCACTTCTCCAGGATCTTACCCGCCGGGAATCTAGTGCC
TTCCGCAGTGAAACCGCCAAAGCCAGAAGATGCTCAAGGAGCTCATCACCAGCACCCGCCTGGGAACCTTACTAC
AACTCCAGCTCCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTTATTCTCCAAATCCCCGAG
CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGCAGGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC
TCGGCTGCCGTCCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
GACATAGCTGCATTGAATTCACGCTGGGTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG
AAGGGCCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACCTCCG
CTGGAGTGGACGCTGGCAGAGTGCAGGGACCGACTGGCCATGTATGACGTGGCCGGGCCCCCTGGAGAAGAGGCTC
ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTCTGGCGTGGGGGGCCATCATGGCGGT
GTCTGGAAGAAGGGCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCTGT
GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTAC
TCGCCCCAAACCACTGCTCCTGGCACCTCACGGTGGCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC
TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGT
GGCTTGCGCATCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC
TCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTGTACAACAGTGGACCCCTGCCCTGGA
GAGTTCCTCTGTTCTGTAATGGAATCTGTGTCCCTGCCTGTGATGGGTCAAGGACTGCCCAACCGCCTGGAT
GAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCCCAAGGTCTGT
GATGGGCAGCCTGATGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCC
TTCAGTGTGAGGACCGGAGCTGCGTGAAGAAGCCCAACCCGCACTGTGATGGGCGGCCCCGACTGCAGGGACGGC
TCGGATGAGGAGCACTGTGACTGTGGCTCCAGGGCCCCCTCCAGCCGCACTGTTGGTGGAGCTGTGTCTCCGAG
GGTGAAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTGACACATCTGTGGGGGGGCCCTCATCGCTGACCGC
TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTTCCTGGGC
AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCAC
GAAGAGGACAGCCATGACTACGACGTGGCGCTGTGTCAGCTCGACACCCGGTGGTGGCTCGGCCCGCGTGGC
CCCGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCTGCACTGCTGGATTACGGGTGGGGCGCCTTG
CGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGCAGCGAG
GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGT
GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTG
GGCTGTGGCCGGCCTAACTACTTCCGGCTTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTG
ACCTGAGGAAGTGGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCCAGGGCAACTGCCAAGCAGG
GGGACAAGTATTCTGGCGGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC
CTGATGTCTGTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGACGTCCCTGAGGACC
CAGGCCACACCCAGCCCTTCTGCCTCCCAATTCTCTCTCCTCCGTCCCCCTTCTTCACTGCTGCCAATGCAAG
GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
CTGTTTGGGCAGCCTTGCTTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCCCTGGTCTAACTTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGAGGGGACCCCTCAGAGCCCTGGAGACTGCCAGGTGGGCCTGCTGCCACTGTAAGCCAA
AAGGTGGGGAAGTCTGACTCCAGGGTCTTGCCCCACCCCTGCCCTGCCACCTGGGCCCTCACAGCCAGACCCCT
CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YYNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGVRLKGPDLASSCLWHLQGPDLML
KLRLLEWTLAECDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLNRLDSQGVLSSTPYFSPSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDCGQPDCLNGSDEEQCQEGVPCGTFTTFQCEDRSCVKKPNPQCDGRPDGRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAAHCFQEDSMASVT
WTVFLGKQVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSSGGLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

[illegible]

紅花

[illegible][illegible]

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSL VNQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

total: 9602001

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCTGGGCAAGGGCCGGGGCGCGGGCCGAGCCACCTCTTCCCCCTCCCCCGC
TTCCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT
CGGCAAAGTTTGGCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGAGCCAGGCCAGACCAGGGGCGCTCG
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGC GGCGGGCGGACGGAGAAAAACAACCTCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG
CCGCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTCCGAGCCCGGGCCATGGAGCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCCGGGGCTCCGCGCGACCCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTCCG
GCACCTCTGGACAGCCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTTAGAAGTGCAGGGCACCTTACAGA
GGCCCCCTGGTCCGGGACAGCCGACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCGGGGGCAACGTACCATCACTTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG
AGTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCTCCCTGCTTGAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCACCCCCCAGT
CTGCCATTGGCTGCTGGACCCCATGATGGCCGGCGGCTGGCCGTTCACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTACCACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCCTTGGGACAGACCCCTGTGGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCCTGCTACCTGCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGCGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGTATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTTCGACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
CTTCCTACGGGCAGCTCATTTGCCAGGGTGCCATCCCACCTGTAGAAGACTTTCCTACAGAGAATCCTAATGATA
ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCCCAGGTG
CCCGCCGTGCTCAGCGGGGCGGCTGATGCGACGCTGGTACGCCGTCTCCGCCGTGGGGCTTGTCCCTCGAA
CCAACACCCCGGCTCGGGCCCTGAGGCCAGATCCAGGTACACCTTCTGCTGCTCCCCTTGGGCCCTAGATG
GTGGCACAGGTCCAGCCCGTGGGGCGGGGAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCA
AGGCTCCCCCTCCCATCTGTAGCAGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGACGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCC
CAGGACCAACCCGAGCCCCCTGGACCCCAACAGCAGTCTTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACTTTTAGAGGTGGGTGAGCCTCCCTCC
ACCACTTCCCTTCCCTGTCCCTGGATTTTCAAGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAACCTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTGGCATCCACAATCCCTCCTACAGGGCCTGG
CTCACAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
GGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCSVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVLTLED FYGVFSSPGYTHLASVSH PQSCHWLLDPHDGRRLAVRFTALDLGFGDAVH
VYDGP GPP ESSRLRLSLTHFSNGKAVTVETLSGQAVVS YHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPCADGSDEWDCS
YVLPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQGRRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLP SLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTTGTAAATTTTATATTACTTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCCAAAAA

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

total = 580200

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC
TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACCTGTGACAGCAGTATGCTGTNTTGCCGAC

Figure 71: Sequence

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCGTGGGCCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCAACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTGCGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTATCTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPIWIVSYALMSVCGAILLVLLVLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGAATTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGACCCGTGAGGCGTTTAACCTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTGCGGGACGTCTTGACTTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTGGATGGCTGCGCCCACTTTTG
CGCCCATTGGCTTGGCTGGTGCTCCGGGCACCAAGAGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAAC TGATCCCTTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCTCGGAATAAAGCGGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGA YVLVYYNLVKAPPCGGMGNLRGRTAVVTGANS
GIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLS
SEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAH
CRGRLD FKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSEL
FLRHVPGWLRPLLR
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARD
DRAAHLWEASKRL
AGLGPGE DAEPDED PQSEDSEAPSSSLSTPHPEEPTVSQYPSPQSSPD
LSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCTGCCTCTGCTGCTCCTGCTCATCATGGGAGGCATGGCTCAG
GACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCACCATCCGCTGGTTGCTGAATGGGCAGCCCTGAGCATGGTGCCCCCAGAC
CCACACACCTCCTGCCTGATGGGACCCCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCACAGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGTGGCTGTCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCACAGTCTCATGGTGGAAAGATGGGAAACCCCTG
GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGTTCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCCATCCAGGAGCCC
CAGGACTACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTGAGCTGGAAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCCAAGCCTAGACCGCGGCTGTGGCTCAGCTGGGAAGGTGAGTGGCCCTGCTGCGCCTGCCCAA
TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCC
GGCTGGCAGAGCGCAGAGCTTGGAGGCCCTCCACTGGGGCCAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCCTGAGGCTGCCGGAAGAGTGCCAGTGCCCCACCTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCCACCACCTGCTGAAAACCAATGGCATC
ATCCGTGGCTACCAAGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAAGTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGGAAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAGTGGCTGCAGTCACTGGTGTGGAGCT
GGGGAGCCAGTAGACCTGTCTGCCCTCTTTAGAGCAGGCCATGGAGCGAGCCACCCAAGAAGCCAGTGAGCAT
GGTCCCTGGACCCTGGAGCAGCTGAGGGCTACCTTGAAGCGGCCTGAGGTGATTGCCACCTGCGGTGTTGCACTC
TGGCTGCTGCTTCTGGGCACCGCCCTGTGTATCCACCGCCGCGCGGAGCTAGGGTGACCTGGGCCAGGTCTG
TACAGATATACAGTGAGGATGCCATCCTAAAAACAGGATGGATCACAGTGACTCCCAGTGGTTGGCAGACACT
TGGCGTTCCACCTTGGCTCTCGGGACCTGAGCAGCAGCAGCCTCAGCAGTGGCTGGGGCGGATGCCCGG
GACCCACTAGACTGTGCTGCTGCTTGGCTCTCCTGGGACTCCCGAAGCCCCGCGGTGCCCTGCTTCCAGACACC
AGCACTTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCAGCCAGGCCAAGTCCCCAGGTCCCAGCT
GTCAGGCGCCTCCACCCAGCTGGCCAGCTCTCCAGCCCTGTTCCAGCTCAGACAGCCTCTGCAGCCGAGG
GGACTCTCTTCTCCCGCTTGTCTCTGGCCCTGTCAGAGGCTTGAAGGCCAAAAGAAGCAGGAGCTGCAGCAT
GCCAACAGTTCCTCACTGCTCCGGGGCAGCCACTCCTTGGAGCTCCGGGCCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCCTTTCCCAAGCCAGGAGCTGTGCCCCAAGCTCTGGTTGCTGGCGGGCCCTGGGACCGAAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCACTCTCCCTCCAGCACCCCTCTTTCTCATGAAACTCCCCAACTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACCACAGGCTCCCTCCTCCATCCTGCTGCCAGCAGCCCCCATCCCCATC
CTTAGCCCCCTGAGTCCCCCTAGCCCCAGGCTCTTCCCTCTCTGGCCCCAGCCAGCTTCCAGTGCCTGTCC
AGCTCCTCACTGTCACTCCCTGGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTAGTGAGGGTGAGGAGACTCCAGGAACAGCGTCTCTCCATGCCAAGGGCTCCTTACCCCCCACCACCTAT
GGGTACATCAGCGTCCCAACAGCCTCAGAGTTACGGACATGGGCAGGACTGGAGGAGGGGTGGGGCCCCAAGGGG
GGAGTCTTGTGTGCCCCACCTCGGCCCTGCCCTACCCCCACCCCCAGCGAGGGCTCCTTAGCCAATGGTTGGGGC
TCAGCCTCTGAGGACAAATGCCGCCAGCGCCAGAGCCAGCCTTGTGAGCTCCTCCGATGGCTCCTTCTCTGCTGAT
GCTCACTTTGCCCGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCAGGGAGGCAGACTGC
GTCTTCATAGATGCCCTCATCACTCCCTCCCCACGGGATGAGATCTTCTGACCCCCAACCTCTCCCTGCCCTTG
TGGAGTGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACACCAGCGGCTGGGAAGGGGGATGCCCTCCC
TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAAGTCACTGCTGATGCCAAGGCTGGTGTCTCTCCT
GTAGATTACTCTGAACCGTGTCCCTGAGACTTCCAGACGGGAATCAGAACCCTTCTCCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
TTCTCCCTCCACGATTGTGAAAAACAAATGAAAAACAAATAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAAA
ACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCTCTGTGCCATCCACTCCCACCACAGGTTGTTTGGC
CTGAGGAGCAGCCCTGCCTGCTGCTTCCCCCACCATTGATCACAGGAAGTGGAGGAGCCAGAGGTGCCCTTT
GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGAGCCCCCTCTCAGCCTTACCT
GGGCCCCCTCCTCTAGAGAAGAGCTCAACTCTCTCCCAACCTCACCATGGAAAGAAAATAATTATGAATGCCACTG
AGGCACTGAGGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAAGCACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAAATTTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGP GPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHHLLP
DGTLLLLQPPARGHAHDGQALSTD LGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPG RHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLE NVTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVA AVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGWP
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLL PDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRLSSPRLSLAPAEA
WKAKKKQELQHANSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCP RPRCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSH TQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCAGAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
ACCTCACCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTCTTTCCAGGACCCTGCTGTCTCCTCCCTCCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGCTGCTGAGCCTTCTGTCAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCCGACTCCAGGGCTCTCCCCACCTCCTCCAGGCTCTCCTCTTGTCATGTTCCA
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCTGGGC
CTCATGCCCAGTGTGCGACCCCTGCCTTCTCCTCCACTCCAGACCCACCTTGTCTTCCCTCCC
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGACG
GACTCTGAATTCTAACAATGCCAGTGAAGTGTGCGACTTGAGTTTGAAGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTTACGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGAATCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCAGGACCGGCACAGAAGTGG
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCCTTCTTGCCCTTTGGAAAAAATGATGAAGA
AAACCTTGCTCCTTCTTGTCTGGAAAGGGTTACTTGCCATATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
GCACAACTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCTCCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCTTTTAAAAAGTGAAATTAAGAGTTGTTTCAATATGCAAACTTGGAAAG
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTCTGTTTTGTGTTACTTCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT
TTTACAGAGCAATTATCTTGTATATACAACCTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTATAAATAAAATGTTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPPEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSIRD SRQELSLIVTLWNLTLDAGEYWCGVEKRGPD E SLLISLFV
FPGPCCPPSPSPSTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSP PAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE E KEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTCAACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTNPVNPVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGCD

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCGGGCGCCCGCGGCGGCACCATGAGTCCCCGCTCGTGCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTCTGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAGTGTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGACAGTGGGGAGCT
GGAGAAGTGCGGCTGTGACAGGACAGTGATGGGGTCAGCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGGGAATGCAAGTGCCACGGGGTGTGAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGCTGCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTTTTGTTCTGGTTTTGGTTTTTGGGTCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCCGATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTCATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAAGAACTCTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATACCAGGTCAAGGGCACCAGGTTTCATTTAGCCCTTACATGGACAGCTAGA
GGTTGATATCTGTGGGTCCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCACCCCTAGAACCCAGCCTGCCCCAGCCTGCCCCTGGGAAGAGGAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTTCTCCATGAAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCCTTTCAGTCATGCCCGAGTCACCTTTCACAGCACTGTTCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPLVFGKVVTQGTREAAFVYAISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAAGTGCCTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACCTCCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTCAAGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGC
TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACCTGCTTGACCCCGCCCAGTGCCCTGGAGCCGCGCCCATTCAGCATGTCGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA
AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVLTQGLAEAGRGREDEVTELFRALAEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENVCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCTTCTCCTCCTTTCTCCACGTCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACCTGGAGCCTCATTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAACCCAGCCCGGCCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTAGTAACGGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTTCGGCGCCCGCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTTCGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCTCTCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCCTGCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCTGCTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCGAGCCCCCTGGGGCCCCCGG
GAGCCATGGGGTGTCGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTTCATCG
TCCAGGGGCCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSLDLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLP SRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTKSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV

Important features:

Signal peptide:

amino acids 1-26

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAATA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGAGAAATTTT
ATCCAACCTTTGTTTGGGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTTCATTCACGGTTGCGTTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTTCGAGCAGATGAACT
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAAATAAAAAGCTGT
GAGGTTTTGTTTAAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCGAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACCTCTAGCAG
CAAAACAAAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAGAAGATAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAATGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCTT
TCCACAAAAA

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919

><subunit 1 of 1, 472 aa, 1 stop

><MW: 53847, pI: 5.75, NX(S/T): 2

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEEVNRSQSMKGKSKSSHDLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKS
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAEYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQDSDTFEIYDPRNPVNKRREESKKLMREKKERR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCTGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAAGTGTCCCAGACCCATTTGCGCTTGCTGACGGCGTCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTTGGAAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTATAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCTCG
ATGTACGTGAGCACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCCG
GAGAAAGTGGTGCTGGTGTTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTTCGGCCAGTCCGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTCAACCCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCAGAGTGGCACCGCGTTATTGAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTTCCAAC
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCTTATAATATCACCAAGGAGCAGGT
ACCCTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCTGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACTGGGGACAAGAGTTCTACCCACCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCCCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTATACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTTTGGTAGTTTGGGA
TCTTCTCTCCACCCACACTTATCTCCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQKGQMHVGKTPIQVFLGVPPFSRPP
LGILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVIYAPARAPGDPQLPVMVWFFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNNKMRFLQLNFQRPDEEIIWSMSPVVDGVVIPPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDKMLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

Year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	
1992	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2
1993	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	15.1	15.2	15.3	15.4	15.5	15.6	15.7	15.8	15.9	16.0	16.1	16.2	16.3	16.4
1994	16.5	16.6	16.7	16.8	16.9	17.0	17.1	17.2	17.3	17.4	17.5	17.6	17.7	17.8	17.9	18.0	18.1	18.2	18.3	18.4	18.5	18.6	18.7	18.8	18.9	19.0	19.1	19.2	19.3	19.4	19.5	19.6
1995	19.7	19.8	19.9	20.0	20.1	20.2	20.3	20.4	20.5	20.6	20.7	20.8	20.9	21.0	21.1	21.2	21.3	21.4	21.5	21.6	21.7	21.8	21.9	22.0	22.1	22.2	22.3	22.4	22.5	22.6	22.7	22.8
1996	22.9	23.0	23.1	23.2	23.3	23.4	23.5	23.6	23.7	23.8	23.9	24.0	24.1	24.2	24.3	24.4	24.5	24.6	24.7	24.8	24.9	25.0	25.1	25.2	25.3	25.4	25.5	25.6	25.7	25.8	25.9	26.0
1997	26.1	26.2	26.3	26.4	26.5	26.6	26.7	26.8	26.9	27.0	27.1	27.2	27.3	27.4	27.5	27.6	27.7	27.8	27.9	28.0	28.1	28.2	28.3	28.4	28.5	28.6	28.7	28.8	28.9	29.0	29.1	29.2
1998	29.3	29.4	29.5	29.6	29.7	29.8	29.9	30.0	30.1	30.2	30.3	30.4	30.5	30.6	30.7	30.8	30.9	31.0	31.1	31.2	31.3	31.4	31.5	31.6	31.7	31.8	31.9	32.0	32.1	32.2	32.3	32.4
1999	32.5	32.6	32.7	32.8	32.9	33.0	33.1	33.2	33.3	33.4	33.5	33.6	33.7	33.8	33.9	34.0	34.1	34.2	34.3	34.4	34.5	34.6	34.7	34.8	34.9	35.0	35.1	35.2	35.3	35.4	35.5	35.6
2000	35.7	35.8	35.9	36.0	36.1	36.2	36.3	36.4	36.5	36.6	36.7	36.8	36.9	37.0	37.1	37.2	37.3	37.4	37.5	37.6	37.7	37.8	37.9	38.0	38.1	38.2	38.3	38.4	38.5	38.6	38.7	38.8
2001	38.9	39.0	39.1	39.2	39.3	39.4	39.5	39.6	39.7	39.8	39.9	40.0	40.1	40.2	40.3	40.4	40.5	40.6	40.7	40.8	40.9	41.0	41.1	41.2	41.3	41.4	41.5	41.6	41.7	41.8	41.9	42.0
2002	42.1	42.2	42.3	42.4	42.5	42.6	42.7	42.8	42.9	43.0	43.1	43.2	43.3	43.4	43.5	43.6	43.7	43.8	43.9	44.0	44.1	44.2	44.3	44.4	44.5	44.6	44.7	44.8	44.9	45.0	45.1	45.2
2003	45.3	45.4	45.5	45.6																												

[illegible]

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPKAGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPRPDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLEFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins
signature.**

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCTCTCTCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCTATCAAAGTTTCGGATATCAATGACAATGAACCAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAACCTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTTCGCAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTAAA
AAAGAAAGTGGATTTTGGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCTTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTTATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCTTCCATATTATGTATTTGAAGTTTGTGA
AGAAACCCACAGGGATCATTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTTGACTAATAGAAGTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTGTCATTATGATCATA
TTTGGGTTTTATTTTTTTGACTTTGGGTTTTAAAACAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAGAGAGAATATATTTCCAATATGATGATGAAGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCCTCTTTTGGATTCCCTCCAGACCTACGCTTTTGGAGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTTTACCATCAAATTTTTTAAAAGTGCTAATGTGTATTGCAACCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTAAAGCTACCTACATGCTGTCTATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAATAACCAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTCCTTTAT
TTAAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVS DINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLOQQPYFSVEPTTG VIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSLSD
VNDNKPIFKESLYRLTVSESAPTGT SIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHFHYFNLSVEDTMNSSFTIIDNQDNTAVILTNRGTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDCGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTGCGAAACATTTGACATTATT

Figure 96

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGGCGGCGCAGCGAGCAGGGCTCCCCGCTTAACTTCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCGAGCGAGTCATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCTGCGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGACAGGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACCACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAACAAAA
ACCCATGTGTTAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCTCA
ATATAGGAGGGAAGATTTTTCATTTGTATTACTGCTTCCCATTTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTCTATTAAAA
ATAGACAGTAAATACTATTCTCATTATGTTGATACTAGCATACTTAAATATCTCTAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCC
TTATATACATATGTAACAGTCAAATATCATTACTCTTCTTCACTAGCTTTGGGTGCCCTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCACTCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATG
ACCAAAGTGATAAATTCCTGTGACCTTCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACTTTATTGATTGAATTTTAAAGCTACTTATTCATAGTTTATATCCCCCTAACT
ACCTTTTGTTCCTTCCCTAATTGTAATTGTTTTCCTCAAGTGTAATTATCATGCGTTTTTA
TATCTTCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACATATGCCTATGTAGTGTCTTTATTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACTTCAT
GTGGTTCAGTGCCCTTCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCCTTCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACAACCTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTTGGTTTTGTAAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACCTGCATACGTTTTGGTGTGTGCTTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI¹LAFLGWIGAI²VSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSL³LNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLE⁴DDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMT⁵PNARYEFGQALFTGWAAASLCLLGALLCCSC
PRKTT⁶SYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

Parameter	Unit	Value	Standard Error	95% CI	P-value
Intercept		1.00	0.00	1.00	0.00
Age	Year	0.02	0.01	-0.01, 0.05	0.10
Sex					
Male		0.05	0.02	-0.01, 0.11	0.08
Female		0.01	0.02	-0.03, 0.05	0.70
Education	Year	0.01	0.01	-0.01, 0.03	0.30
Income	Year	0.01	0.01	-0.01, 0.03	0.30
Health status					
Good		0.05	0.02	-0.01, 0.11	0.08
Poor		0.01	0.02	-0.03, 0.05	0.70
Smoking status					
Smoker		0.05	0.02	-0.01, 0.11	0.08
Nonsmoker		0.01	0.02	-0.03, 0.05	0.70
Alcohol consumption					
Drinker		0.05	0.02	-0.01, 0.11	0.08
Nondrinker		0.01	0.02	-0.03, 0.05	0.70
Physical activity					
Active		0.05	0.02	-0.01, 0.11	0.08
Inactive		0.01	0.02	-0.03, 0.05	0.70
Stress					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Depression					
Yes		0.05	0.02	-0.01, 0.11	0.08
No		0.01	0.02	-0.03, 0.05	0.70
Family size					
Large		0.05	0.02	-0.01, 0.11	0.08
Small		0.01	0.02	-0.03, 0.05	0.70
Marital status					
Married		0.05	0.02	-0.01, 0.11	0.08
Single		0.01	0.02	-0.03, 0.05	0.70
Religious belief					
Religious		0.05	0.02	-0.01, 0.11	0.08
Non-religious		0.01	0.02	-0.03, 0.05	0.70
Work status					
Employed		0.05	0.02	-0.01, 0.11	0.08
Unemployed		0.01	0.02	-0.03, 0.05	0.70
Living alone					
Yes		0.05	0.02	-0.01, 0.11	0.08
No		0.01	0.02	-0.03, 0.05	0.70
Health insurance					
Yes		0.05	0.02	-0.01, 0.11	0.08
No		0.01	0.02	-0.03, 0.05	0.70
Healthcare access					
Good		0.05	0.02	-0.01, 0.11	0.08
Poor		0.01	0.02	-0.03, 0.05	0.70
Healthcare utilization					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Healthcare satisfaction					
Satisfied		0.05	0.02	-0.01, 0.11	0.08
Dissatisfied		0.01	0.02	-0.03, 0.05	0.70
Healthcare cost					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Healthcare quality					
Good		0.05	0.02	-0.01, 0.11	0.08
Poor		0.01	0.02	-0.03, 0.05	0.70
Healthcare equity					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Healthcare sustainability					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Healthcare effectiveness					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Healthcare efficiency					
High		0.05	0.02	-0.01	

Parameter	Unit	Value	Standard Error	95% CI	P-value
Intercept		1.00	0.00	1.00	0.00
Age	Year	0.02	0.01	-0.01, 0.05	0.10
Sex					
Male		0.05	0.02	-0.01, 0.11	0.08
Female		0.01	0.02	-0.03, 0.05	0.70
Education	Year	0.01	0.01	-0.01, 0.03	0.30
Income	Year	0.01	0.01	-0.01, 0.03	0.30
Health status					
Good		0.05	0.02	-0.01, 0.11	0.08
Poor		0.01	0.02	-0.03, 0.05	0.70
Smoking status					
Smoker		0.05	0.02	-0.01, 0.11	0.08
Nonsmoker		0.01	0.02	-0.03, 0.05	0.70
Alcohol consumption					
Drinker		0.05	0.02	-0.01, 0.11	0.08
Nondrinker		0.01	0.02	-0.03, 0.05	0.70
Physical activity					
Active		0.05	0.02	-0.01, 0.11	0.08
Inactive		0.01	0.02	-0.03, 0.05	0.70
Stress					
High		0.05	0.02	-0.01, 0.11	0.08
Low		0.01	0.02	-0.03, 0.05	0.70
Depression					
Yes		0.05	0.02	-0.01, 0.11	0.08
No		0.01	0.02	-0.03, 0.05	0.70
Family size					
Large		0.05	0.02	-0.01, 0.11	0.08
Small		0.01	0.02	-0.03, 0.05	0.70
Marital status					
Married		0.05	0.02	-0.01, 0.11	0.08
Single		0.01	0.02	-0.03, 0.05	0.70
Religion					
Christian		0.05	0.02	-0.01, 0.11	0.08
Muslim		0.01	0.02	-0.03, 0.05	0.70
Other		0.01	0.02	-0.03, 0.05	0.70
Region					
Urban		0.05	0.02	-0.01, 0.11	0.08
Rural		0.01	0.02	-0.03, 0.05	0.70
Season					
Spring		0.05	0.02	-0.01, 0.11	0.08
Summer		0.01	0.02	-0.03, 0.05	0.70
Autumn		0.01	0.02	-0.03, 0.05	0.70
Winter		0.01	0.02	-0.03, 0.05	0.70
Time of day					
Morning		0.05	0.02	-0.01, 0.11	0.08
Afternoon		0.01	0.02	-0.03, 0.05	0.70
Evening		0.01	0.02	-0.03, 0.05	0.70
Month					
January		0.05	0.02	-0.01, 0.11	0.08
February		0.01	0.02	-0.03, 0.05	0.70
March		0.01	0.02	-0.03, 0.05	0.70
April		0.01	0.02	-0.03, 0.05	0.70
May		0.01	0.02	-0.03, 0.05	0.70
June		0.01	0.02	-0.03, 0.05	0.70
July		0.01	0.02	-0.03, 0.05	0.70
August		0.01	0.02	-0.03, 0.05	0.70
September		0.01	0.02	-0.03, 0.05	0.70
October		0.01	0.02	-0.03, 0.05	0.70
November		0.01	0.02	-0.03, 0.05	0.70
December		0.01	0.02	-0.03, 0.05	0.70
Day of the week					
Monday		0.05	0.02	-0.01, 0.11	0.08
Tuesday		0.01	0.02	-0.03, 0.05	0.70
Wednesday		0.01	0.02	-0.03, 0.05	0.70
Thursday		0.01	0.02	-0.03, 0.05	0.70
Friday		0.01	0.02	-0.03, 0.05	0.70

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
 GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCATCNTCAGCACTGCCCTGCCCCAGTG
 GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCAGGCCNTGTACGAGGGGGCTGT
 GGATGTCCTGCGTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
 GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
 TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
 GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCCAGGTCTGGCTA
 TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAAGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTGCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCCTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTCTGTCCTCCCGAA

Figure 103

Figure 1 displays a 12x12 grid of 144 small plots, each showing the evolution of the 1000-bp window of the 1000 Genomes Project. The columns are labeled with dates from 20100101 to 20120101. The rows are labeled with genomic coordinates from 100000000 to 1000000000. Each plot shows a line graph of the number of variants (y-axis, 0 to 1000) over time (x-axis, 0 to 1000). The plots show a general increase in the number of variants over time, with some plots showing a sharp increase around 20110101.

[illegible]

FIGURE 105

TCATAGGGGGGCGCGATATTTTTTCTTG CAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGT CAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCCTG

[illegible]

FIGURE 106

TTCCTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGACAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCCAAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGACGAGGGGCTGTGGATGTCNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCCTATGACCCCAAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTTCTGCGAA

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCGCGGTGCCCCG
ACCGGTCCCCGCTTTTTGTAAACCTTAAAGCGGGCGCAGCATTAAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCAGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTTCAGTCTATGTTTGCTCCAAGTACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACCTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAATAAAATTATATCCACAAGTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAAACAAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTTCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTAAGTGCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGAGTGAAGTCAAAGTCTTATTGAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAAGTCTGTTGGGTGAAGTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTGAGTAAATAACAAGTGTATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTLNKLGNPTDRNVCFKVKTAPRRYCVRPNSGIID
AGASINVSVMQLQPFDDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLMRKTVQSNISPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

Figure 110

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGC GGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTA CTGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAAC TGTATTTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

Figure 111

FIGURE 112

CCCTGGTGGTTTTGTTCCTTAATTGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

Parameter	Value	Unit
α	0.000000	
β	0.000000	
γ	0.000000	
δ	0.000000	
ϵ	0.000000	
ζ	0.000000	
η	0.000000	
θ	0.000000	
ι	0.000000	
κ	0.000000	
λ	0.000000	
μ	0.000000	
ν	0.000000	
ξ	0.000000	
π	0.000000	
ρ	0.000000	
σ	0.000000	
τ	0.000000	
υ	0.000000	
ϕ	0.000000	
χ	0.000000	
ψ	0.000000	
ω	0.000000	
Ω	0.000000	
Θ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	
Φ	0.000000	
Ψ	0.000000	
Ξ	0.000000	
Π	0.000000	
Σ	0.000000	
Υ	0.000000	

FIGURE 113

GGTGGCCCATTCCTCGGCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTTANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTG
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

GGTGGCCCATTCCTCGGCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCGTGCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACGTATTATTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

[illegible]

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANTCCC
GGCCCAGGCTGCTTTCCGTGTCCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 115

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGGCCAGGCTGCTTTCGGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTT
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAATAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCAGTTGAGAAA
GGTACCCCGGATTGTGAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTCTGAATTGGAGGATTATCTTTTCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTTCAAGCATCTTGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCACGGGGCTGTAGTGGCAT
TCTCATTTCCCTCAGCATGTTCTAACTGCTGCCACTGTGTTTATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTTC
AGAGGATTGCCGAAGGGAGGCCCTTCTTTTCAAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTTCAAGAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCCATAAAATACGCCCAGATTTGCCTCTGGATTACGG
GAACGATGCCAATTGTGCTTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACCTCTGTCAATAGCATTTCACATTTTTTCAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACATAATACTGTTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFAEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAACVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDSNLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNAVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCCGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCAACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACCTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTGAGGAGGCCCTCGTAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTC SNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPSLKCQ
AKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHRLVLKGPDLHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFP
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQELEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

protein: seqdb2

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAAC TCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTCTGTTGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAACTGGCTAACATTCTTTT
TACCAGGGAAGTACCCCGCCGCTTAGAAGGCACAAATGTACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTGATGGGCTTTTTTCAAACTCCAGTAGAAGGTGCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTACTGAAAAATTATTTT
TGGGATAAGAGAATTTAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAATAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTGAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEAAAGQLRREL RQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNH LGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSLANILFTRELARRLEG TNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTTGGGATAAGAGAATTTTCAGCAAAGATGTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTGAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

Figure 123

[illegible]

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCAATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACCTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTATGCTGACTGGTGT
CGTTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTTCAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTTACATGAAAGAAGATACAGAAAGTTTAGAAATATTCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAACTCCAGCAGATTGTCTCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCTTGGAAGAACTCAAGCAATTCTG
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACCTGAT
ACAGCCCCAGGAGAGCAAGCCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQOKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

protein = 5306001

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

Project: S8041001

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTTCGCCGCNCGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGCTCCCGGAGCCCAGCCCTTTCTTAACCC
AACCCAACCTAGCCCNGTCCCGAGCCGCCAGCGCTGTCCCTGTCNCGGANCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTTACTCCTGTAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCACGCGTCCGATGGCGTTACGTTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTA TACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG
TTTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAACCTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACCTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTTCTTTAGT
GCAATACAATAAAACTCTGAAATTAAGACTC

protein: 5802601

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

protein = 9302001

FIGURE 130

ATTATAGCATTGTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTTATCTTCTAGCATTTTTT
TACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

[illegible]

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

CGGACGCGTGGGGGAAACCCTTCGCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGGG
GAACAAGATGCGCGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAATTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTTAAGATGCCTCTCTCTTAACCTCTGGGTGGATTTTAACTACAACCTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACCTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCCTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSILWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLPLTLVRSFWSMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTTCGC
TGAAGTGAAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

Figure 133

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

total = 530,000

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGGCCGCGGCTG
GGGATTCTTGTTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAATAATGGAATGTCATCTACGAAGAAAAGTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG
CATTTTGATGAGAATTCATTTTTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAA
AAAAAAAAAAAAAAAAAAAAA

[illegible]

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLSESDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIIVSLFNAFGRISTSVKELENFRNLLONIH

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCAT
CTTNGAGCGCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTCCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

Sequence

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGTATGGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAATAATTTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTACAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 138

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCVPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

protein sequence

FIGURE 140

CACAGTTCCCCCACCATCACTCNTCCCATTCCTTCCAACCTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCTTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGGCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

Figure 1 consists of 12 histograms arranged in a 6x2 grid. The left column contains histograms for $n = 10, 20, 30, 40, 50, 60$ and the right column contains histograms for $n = 70, 80, 90, 100, 110, 120$. Each histogram plots the 'Number of non-zero elements' on the x-axis against the 'Frequency' on the y-axis. The distributions are approximately bell-shaped and centered at $n/2$. As n increases, the frequency of the central peak increases significantly.

Parameter	Value	Standard Error	z	P	95% CI
Intercept	1.00	0.00			
Age	0.01	0.01	0.10	0.92	-0.01, 0.03
Gender	0.01	0.01	0.10	0.92	-0.01, 0.03
Education	0.01	0.01	0.10	0.92	-0.01, 0.03
Income	0.01	0.01	0.10	0.92	-0.01, 0.03
Health	0.01	0.01	0.10	0.92	-0.01, 0.03
Marital	0.01	0.01	0.10	0.92	-0.01, 0.03
Religious	0.01	0.01	0.10	0.92	-0.01, 0.03
Political	0.01	0.01	0.10	0.92	-0.01, 0.03
Occupational	0.01	0.01	0.10	0.92	-0.01, 0.03
Family	0.01	0.01	0.10	0.92	-0.01, 0.03
Community	0.01	0.01	0.10	0.92	-0.01, 0.03
National	0.01	0.01	0.10	0.92	-0.01, 0.03
Global	0.01	0.01	0.10	0.92	-0.01, 0.03
World	0.01	0.01	0.10	0.92	-0.01, 0.03
Universe	0.01	0.01	0.10	0.92	-0.01, 0.03
Population	0.01	0.01	0.10	0.92	-0.01, 0.03
Society	0.01	0.01	0.10	0.92	-0.01, 0.03
Culture	0.01	0.01	0.10	0.92	-0.01, 0.03
Environment	0.01	0.01	0.10	0.92	-0.01, 0.03
System	0.01	0.01	0.10	0.92	-0.01, 0.03
Structure	0.01	0.01	0.10	0.92	-0.01, 0.03
Organization	0.01	0.01	0.10	0.92	-0.01, 0.03
Arrangement	0.01	0.01	0.10	0.92	-0.01, 0.03
Configuration	0.01	0.01	0.10	0.92	-0.01, 0.03
Design	0.01	0.01	0.10	0.92	-0.01, 0.03
Pattern	0.01	0.01	0.10	0.92	-0.01, 0.03
Form	0.01	0.01	0.10	0.92	-0.01, 0.03
Shape	0.01	0.01	0.10	0.92	-0.01, 0.03
Figure	0.01	0.01	0.10	0.92	-0.01, 0.03
Image	0.01	0.01	0.10	0.92	-0.01, 0.03
Picture	0.01	0.01	0.10	0.92	-0.01, 0.03
View	0.01	0.01	0.10	0.92	-0.01, 0.03
Scene	0.01	0.01	0.10	0.92	-0.01, 0.03
Environment	0.01	0.01	0.10	0.92	-0.01, 0.03
World	0.01	0.01	0.10	0.92	-0.01, 0.03
Universe	0.01	0.01	0.10	0.92	-0.01, 0.03
Population	0.01	0.01	0.10	0.92	-0.01, 0.03
Society	0.01	0.01	0.10	0.92	-0.01, 0.03
Culture	0.01	0.01	0.10	0.92	-0.01, 0.03
Environment	0.01	0.01	0.10	0.92	-0.01, 0.03
System	0.01	0.01	0.10	0.92	-0.01, 0.03
Structure	0.01	0.01	0.10	0.92	-0.01, 0.03
Organization	0.01	0.01	0.10	0.92	-0.01, 0.03
Arrangement	0.01	0.01	0.10	0.92	-0.01, 0.03
Configuration	0.01	0.01	0.10	0.92	-0.01, 0.03
Design	0.01	0.01	0.10	0.92	-0.01, 0.03
Pattern	0.01	0.01	0.10	0.92	-0.01, 0.03
Form	0.01	0.01	0.10	0.92	-0.01, 0.03
Shape	0.01	0.01	0.10	0.92	-0.01, 0.03
Figure	0.01	0.01	0.10	0.92	-0.01, 0.03
Image	0.01	0.01	0.10	0.92	-0.01, 0.03
Picture	0.01	0.01	0.10	0.92	-0.01, 0.03
View	0.01	0.01	0.10	0.92	-0.01, 0.03
Scene	0.01	0.01	0.10	0.92	-0.01, 0.03
Environment	0.01	0.01	0.10	0.92	-0.01, 0.03
World	0.01	0.01	0.10	0.92	-0.01, 0.03
Universe	0.01	0.01	0.10	0.92	

AAGATTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
 GGCAGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
 ATTCAAGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
 GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
 AATGCAGACTTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTTCATGTGGTTTTTCT
 ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
 TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCTTGGAGA
 AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
 GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
 ACGGCCACTGTGCCATACAACCTTCGTGTCTAGGGCCACATTGGGCTCACAGACCTCAGCCTG
 GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
 TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
 CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
 GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
 CATTCTGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
 GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCCTTTGTTGGCTTCATGCTGATCCTTGT
 GGTCTGTGCCACTGTTCTGTCTGAAAATGGGCCCGCTGCTCCAGTACTCCTGTTGCCCCGTGG
 TGGTCTTCCCAGACACCTTGAAAATAACCAATTACCCCCAGAAGTTAATCAGCTGCAGAAGG
 GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
 CTCATAGGTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGA AAC
 ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
 GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACAC
 TGA CTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
 CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
 TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
 TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATT CAG
 TGTTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
 GGCTTGAGAGAGCCCACTTTCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
 TGTTGAGTTCACTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
 AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
 CATGAACCTACTGTAAAGTGTGTGACAGTGTGTGCACTGCAGACAGCAGGTGAAATGTATGT
 GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
 TGTTGGTAAAGTACAGAATT CAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAA
 AAAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLSFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHELLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS AW
SILKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAE EHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLV LALFAFVGFM LILV
VVPLFVWKMGRL LQYSCCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

protein sequence

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCCTCCTTTACAGCCACCGACTCAGTGAAGTGGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATCTACTTCAACCAGGAACCTCTACGGGAATTTTCAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCTATTCCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTCT
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTTCGCTGAGACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLFLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLRDITITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSLTPPCSE
TVTWILIDRALNITSLQMHSRLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPR
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 145

FIGURE 146

GGCGCCTGGTTCCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC
 GTTCGTGTCCCCGCCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA
 GCAGCTCCCTTCCACCCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCCTGCCATTTCTTCCAGGTTGAGGGAGC
 CGCAGAGGCGGAGGCTCGCGTATTCTTGCAGTCAGCACCCACGTGCGCCCCCGACGCTCGGTGCTCAGGCCCTTC
 GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
 CACCTCTCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
 ATCCTGAGGTCATTTCATTATGAAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGCTGTTTCATGGCTAGA
 GCAATTCCAGCCATGGTGGTTCCTCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
 GAGTGGTGGATAGCCAAACAACGAGGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
 AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
 TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
 TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA
 CAGGTGCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCTAATCATGAACATCTGGGGGCGAG
 ATATGGCCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGGAACTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGCCCTGTTCTGCTTGCCACCTAGTTTTTGGAGGGGGCTGTAGAGAAAATCTGTGTACAAAGAAGGG
 TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
 CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT
 TGTGAAGTAAGATTAAAGAGATCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCTGTGGCTGTTTGGAT
 AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
 ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA
 AATGGTATTCAAACAATTGGCAATATCAGTCTGCTAATTCCTTTCACAGTCTCTAAAGTAACAGTTTCAGGCTGTG
 ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTTCACATTGCCCAAGAGTATACTGTCTCT
 CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATTGGAACCTCGAGTTTATTCTGATCTGTCCAGTATC
 TGCAGAGCAGCAGTACATGCTGGAGTGGTTGCAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
 AAGACCTACATTGCTTCTTTTTCAGAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATTC
 AGAGTGTTTGTGCTGTTGTGTAAGTGAATCTTGGAAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCCAAAAGGTGCCAAATGCATA
 TAAATCTTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTTGGGAAAAGTAATGAAAATATAATGG
 TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
 GTTCTACGTTTTCATATATTATATGGTGTCTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
 GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCC
 AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTCT
 TGTTAATTTAGGCATATAGAATATTAAATTTCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC
 CAAATGAATCTGTTAAATGTTTGATTCCCTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
 ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
 AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
 GAACAAAGATGAACATAATGTATTACATTACCATTTGCCACTGATTTTTTTTTTAAATGGTAAATGACCTTGTATATAA
 ATATTGCCATATCATGGTACCTATAATGGTGTATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
 AATCTGTAAATGTTAGTTTTTGGTAATTTTTTTCTGCTGGGATTTACATATTAAATTTTTTCTGCTGGTGGAA
 TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLEKYMDEDEGEWWIAKQRGKRAITDNDM
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQV VWATSNRIGCAINLC
HNMNIWGQIWP KAVYLV CNYS PKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YEC PAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSES LQNPPGGKAFRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

60407:330400

2000 10000 20000 30000 40000 50000 60000 70000 80000 90000 100000
 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
 CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
 GGCGTCTCCGGGCCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCACGCCC
 CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
 ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
 AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACCTGCGAGGAGAAGATGGTTAT
 CATCACCAACAAGAGCGTGTCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
 AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
 GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
 CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
 TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTACCACGGTCAG
 TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCAGATGGGAGACCCATCTCTCTTGCTGCT
 CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTTAA
 AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
 GGAACAATGAGCTTGGTGACACATTTCAATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
 TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
 GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCGTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
 CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATCCCCCT
 GGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
 CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
 TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTGAGCAAAAACCTTAGGAGAAAACT
 TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
 AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
 TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
 GACCATGCGCCCTTTTCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
 CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTGAGAAACCATTAGCAT
 TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
 AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMWIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

protein = 330.400

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTACCCCGGTCTGCGTCATGTTAAACTCCAATGTCTCTCTGTG
GTTAACTGCTCTTGCCATCAAGTTACCCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGCCTAAGAACCCTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTA
TGCCCTACCCCCCACTGGAGAGAGGCGGTTTTAGCCCCCAGAACCCCCGTCTCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTTAAACATCTACGT
GCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCA
CATGATTGACGGCAGCATTTTGGCAAGCTACGGAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
AGGGTTTTTAAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAAATGTGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTGACCGCTGTTGACCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTAGAGCGGCACCGC
CCTGTCCAGCTGGGCAGTGAACTACAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTGCGGTGCAACAT
GCTGGACACCACGCACTGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCCCGGTGATCGACGGCGACGTCATCCAGACGACCCCGAGATCCTGATGGA
CAAGGCGAGTTTCTCACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCTGAAGTTCGTGGACGGCAT
CGTGGATAACGAGGACGCTGTGACGCCCCAACGACTTTGACTTCTCCGTGTCCAATTCTGTGGACAACCTTTACGG
CTACCCTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAAACC
GGAGACGCGCGCGGAAAACCTGTTGGCTCTCTTTACTGACCACAGTGGGTGGCCCCCGCGTGGCCGCGGACCT
GCACGCGCAGTACGGCTCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCGAGT
GGCAGATTCCGGCCCATGGTGATGAGGTCCCCTATGTCTTCCGATCCCCATGATCGGTCCACCGAGCTCTTACG
TTGTAACTTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTCATGACCTACTGGACGAACCTCGCCAAAACCTGG
TGATCCAAATCAACCAGTTTCTCAGGATACCAAGTTCATTACACAAAACCCAAACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGTTTGAACCTCGTTTCTCATTTGCACAACTTGAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCTTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
AACCACAAACGCCCAGCAATCACTCCTGCCAACAAATCCCAAACACTCTAAGGACCCTCAAAAACAGGGCCCTGA
GGACACAACCTGTCTCATTTGAAACCAAACGAGATTATTCACCGAATTAAGTGTACCATTGCCGTGCGGGCGTCTC
GCTCCTCTTCTCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGTCTCATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTGTGCTGCAGGCACACGACACTGAGGCTCACCTGCCCGCCAGA
CTACACCCTCAGCTGCGCCGCTCGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATTGATTCAAA
CACACTGACGGGGATGCAGCCTTTGCACACTTTTAAACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTTCCACCCTAGAGTATAGCTTTTGCCCTATTTCCCTTCCCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAGAAAGTCTCCAGACCAGGAATGTTTTGTCCCCT
GACTTAAGACAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCTTATCGTTGGTGTTCCTCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTAAAAATACTGCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAAGGCCCGGGTGTTCCTAACGTATGGAAGCAGCT
GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
CACACAATGGATGGCTCTCTTAAGTGAAGAAAGAGTCAATGAGATTTTGCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAGAAGTGGACTGTGCAGCGAAATCTGTACGGTCTGTGTCAAAGAG
GTGTTTTGCCAGCCTGAACATATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCMNL
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKP SWADSAHGDEV PYV
FGIPMIGPTELFSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPVQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMS
FPYGTRRSPAKIWPTTKRPAITPANNPKH SKDPHKTGPEDTTVLIETKRDYSTE LSVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLT CPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTCTGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCTCTACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTGGTGATTCTGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAAGTGCAGTGGAGAGAACCCAGAAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCTCATCGTCTTTTTCTCCCTGGTGTCTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTTGTGACTGTCAACCATGAGGTATGGAAGGAGCAGGCCTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGAAGTCAAGAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCTGTTGGGTT
TTGCATTTGACCCAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACT
TCTTCCCTGCCTTACCTTCTTTCACTCCATTCAATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTCTGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCAATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQ GALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVD TYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKL FELTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

[illegible]

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPPSVAERSMEGHAPHHFKLVSVHVFI RHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNOWSADQLYLETTGKSRTLQSGLLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSM LCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSHVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLNLVRFVVKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

protein = 330.00

FIGURE 156

AAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
GCCGCTGTTTACCAATCGGGGAGAGAAAAAGCGGAGATCCTGCTCGCCTTGACGCGCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAACAACCGGAGGAGCTCTGACTTCCCACTGTCCCATTCTAT
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT
TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACGCTATTCACTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCGAGCTTTTCGCCCT
GAATCCGCGCAGCGGCAGCTTGGTACGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
TCAATTAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTTCTGTAAGTGAATTAGAAATAAAAATTAGTGAAGTGCAGCCACTGAGATGCGGTT
CCCTCTACCCACGCTGGGATCCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGAACACTCA
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCGAAATTGGTGCTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACCTGGTCCCTACGGCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGCG
CATCCGCGTGATGGTTCTGGATCGGAACGACAACGACCGAGCTTGTCTCAGCCCGAGTACCGCGGAGCGTTCC
GGAGAATCTGGCTTGGGCAAGCAGCTGCTTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCAAGTTTCAAAGTAGATTGTAATTGAGGACAATATC
AACAATAGGGGAGTTGGACCACGAGGAGTCAAGATTCTACCAGATGGAAGTCAAGCAATGGATAATGCAGGATA
TTCTGCGCGAGCCAAAGTCCCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCCCTCACCTCTCT
CGCCAGCTCGGTTCCCGAAAATCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTTCACTCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA
TAGTTTAGTCAAGACATAGTCTTGGATAGGGAAACAGGTTCTTAGCTACAACATCACAGTGACCGCCACTGACCG
GGGAACCCCGCCCTATCCACGGAAGTCAATCTCGCTGAACGTGGCAGACACCAACGACAACCCCGCCGGTCTT
CCCTCAGGCCTCCTATTCCGCTTATATCCAGAGAACAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
CGACCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
GTCTTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
CTTGCAAGTGAAAGTGTGGCGCGGACAAACGGGACCCCGCCCTCAGCAGCAACGTGTCTGTTGAGCCTGTTCTG
GCTGGACCAAGAACGACAATGCGCCCCGAGATCCTGTACCCCGCCCTCCCGACGGACGGTTCCACTGGCGTGGAGCT
GGCTCCCCGCTCCGCGAGAGCCCGGTACCTGGTGAACAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
CTGGCTGTCTACCGTCTGCTCAAGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
CACGGCGGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCACGGCCAGCC
CCCTCTCTCCGCACTGTACGCTCACCGTGGCCGTGGCCGACAGCATCCCCAAGTCCCTGGCGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCCGGCTCTCCTG
CGTCTTCTGGCCTTCGTCTCTTGTCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCTGCTGCAGGC
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCTCGCACTTTGTGGGCGTGGACGGGGTGCAGGCTTTCTGCAGAC
CTATTCCCAGAGTTTCCCTCACACGGAAGTCTCGCGGAAGAGTCACTGATCTTCCCCAGCCCAACTATGCAGA
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAGCGAGCCCTTTTGTGTGAGGTGATTGCGTATTTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTTTTTTTTTAAATTGCTCTGTCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAACCTCCTAGGCTCAAGCAATTATCCCACTTTGCCT
CCGGTGTAAACAGGAGTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT
CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGGAGACAGCCTGACCAACATGGAGAAACCCGCTCTATACTAA
AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCAGCTACTTGGGAGGCTGAGTCAGGAGAAAT
TGCTTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTCCATTCGACTCCAGCCTGGGCAACAAGAGTG
AACTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVP EELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFP LPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLSVSTAHD PDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVL YALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA
NSETSDLTL YLVVAVAAVSCVFLAFVILL LALRLRRWHKSRL LQASGGGLTGAPASHFVGVD
GVQAF LQTY SHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNC SVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRD YRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCTATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTTCACTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSI PHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENF PDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

Figure 1 consists of 12 horizontal bar charts, one for each month from May to April. Each chart displays the percentage of total catch for six fish species: Yellow perch, Rock bass, White perch, Striped bass, Rock bass, White perch, Striped bass, Rock bass, White perch, Striped bass, Rock bass, and White perch. The x-axis represents the percentage of total catch, ranging from 0 to 100. The y-axis lists the species. The charts show seasonal variations in catch percentages for each species.

Month	Yellow perch	Rock bass	White perch	Striped bass	Rock bass	White perch	Striped bass	Rock bass	White perch	Striped bass	Rock bass	White perch
May	10	20	30	40	50	60	70	80	90	100	110	120
June	15	25	35	45	55	65	75	85	95	105	115	125
July	20	30	40	50	60	70	80	90	100	110	120	130
August	25	35	45	55	65	75	85	95	105	115	125	135
September	30	40	50	60	70	80	90	100	110	120	130	140
October	35	45	55	65	75	85	95	105	115	125	135	145
November	40	50	60	70	80	90	100	110	120	130	140	150
December	45	55	65	75	85	95	105	115	125	135	145	155
January	50	60	70	80	90	100	110	120	130	140	150	160
February	55	65	75	85	95	105	115	125	135	145	155	165
March	60	70	80	90	100	110	120	130	140	150	160	170
April	65	75	85	95	105	115	125	135	145	155	165	175

GGGCGCCGGTGACACCGGCGGGCTGAGCGCCTCTGCGGCCCGGCCCTGCGCGCCCCGGGCCCGC
CGCGCCGCCACGCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCGCGCCCCG
GCCGCGCGCCAGGTGAGCGCTCCGCCCGCGCGAGGCCCGCCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGCGGATTCTCTGCGCGTCAAACCACCTGATCCATAAAAC
ATTTCATCTCTCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCGGGGCCGGGCCGTAGCGGCGGGCGCCTGGA
TGCAGACCCGGCCGCGGGGAGACGGGCGCCCCGCCCGAAACGACTTTTCAGTCCCCGACGCGC
CCCCCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC
GGCTGCCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCCGCGCTACACACGCTGCACCTGGACCGC
TGCGGCCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACC GCGTGGCCCATGTGCACCCGCGATGCCTT
CCGTGACCTTGGCCGCGCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCTTGCTGGCTGGCCGTGACCTCAAACGCTTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACGAGCTTCCGCGAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTACCCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCGAGGGCTCCGAGCCACAGGGTT
CCCCACCTCGGGCCTCGCCCGAGGCCAGGCTGTTACCGCAAGAACC GACCCGCGAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGCTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGAACCC CAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCTCTCCTGATGGACGCCTGCCGCCCGCCACCCCATCTCCACCCCATCATGTTTACAGGG
TTCCGCGGCAGCGTTTGTTCAGAAACGCGCCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLQELGPGFLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLHQN RVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCD CRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGS GALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

[illegible]

CCTGGTTCCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRQVLLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFFPLKNAEDLDVGQNNIENYIIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVTGNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFLGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAA
ADLLTVYLVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPFPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATAACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCCTTATTCTTTTAAATTAGTTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAACTTCAGGCAGATTAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCA¹ACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAAGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCACCCGAATCTTGTAAGAAATATTCAAACATAA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGP GSLLLLFLASHCC LGSARG LFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLD ETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSE RDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQG GELVITSVKRWQKGQREFKRISR SIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGGC
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGGAACCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACCCCCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAACTTCCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTTCATCCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCTGA
GTGTTTCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTTGGGGCCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQ RATANMLATMSNPLGVLVANVLS PVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAA SSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDR TK
HFTEATKIGLC L FSLACVPFALVSQ LQGQTLALAATCSLLGLFGFSVGPVAMELA VEC SFPV
GEGAATGMIFVLGQAE GILIMLAMTALT VRRSEPSLSTCQQGEDPLDWT V S L L L M A G L C T F F
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGV L G P S T A T P E C T A R G
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

Year	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

GTCCACATCCTGCTCAACTGGGTGAGTCCCTCTTAGACCAGCTCTTGTTCATCTTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCCCTGGCAATTCTTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGGCCTTGGGGTCTGCTTGTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTTATTTCTTTAGGGATTTGTCAAGGAGTGACCATCTCAGGTTGAAATACCAAGTGTGAGGGAAGTGCC
ATCTGGTACAGTGTATCGGGAAGCTGTCCAGGAATGGGCCGGGAGGAGGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTAGGTGGACTCTGAGGAAGGCTTGTCTCAGCACAGGCGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCTGTTTCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGG
GCTGGAATCTCTGAGAGCGCCTCTCTGCGAAGCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCC
TAACAACTGCAACCTACATCTGTCTCCAGTGAGCACTTTGCTTGGATGTCATTGTGGGCCCTGATGAGAC
CAAACATGCAACTCATAGTGTGAAGAGCTGGACAGGAAATCCATTGATTTTGTATCTGATGTTAACTGCT
CTATGACAATGGGAACCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCGATGACAATG
CCCTGCGTTTGTCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT
GACCGCCACAGACCTTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCCCTCCAGAGGT
GCTGGACACCTTCAATTTGATGTCGAAGACAGGCCAGGTCATTCTGCGTCGACCTTAGACTATGAAAAGAACCC
TGCCTACGAGTGTGATGTTTCAGGCAAGGGACCTGGTCCCAATCTATCCAGCCATTGCAAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTGCATGAGGCTCCAGCCATCTGGTGTGCAAGC
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCTATGGCAGATGACTTGGATTAGGACACAATGGTTTGTGCTCAAGC
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGCCCCAAATATACCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATCTCAGTGCATCAACGACAATGCACCTGTGTTTGAGAAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAAACAACCTTACCTCTCTTCACTTATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
TAATGAAAAAGTCTCATAACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCATGAACTATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTGAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCTGTGCTCAGCGATGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCAGCGGCCTGACACCACTCCACTGCGCCACTCAGCTCCCCGCCATT
CCTTTTGACAACCATTTGTGGCAAGAGATGCAGACTCGGGGCCAAATGGAGAGCCCTTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTCATCTCAACCCTCATAACGGGGCAGCTGTTTGTCAATGTCAACCAATGCCACAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCCTTACAGACCCGAGCCCTGTT
GAGGGTCATGTTTGTCAACAGTGTGGACACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTGAT
GCTGACGGTGATCTGCCTGGCTGACTGTTGGGCATCTTCGGGTGATCTGGCTTGTGTTTGTGATGTCATCTGCCG
GACAGAAAAGAAGGACAACAGGGGCTACTCAACTGTTCGGGAGGCGAGTCCACCTACCCGAGCAGCCCAAGAGGCC
CCAGAAACACATTGAGAAGGCAGACATCCACCTCGTGCTGTGCTCAGGGTTCAGGAGGTCAGGAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCAAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCTGACAGGCCCTT
CCACCTCACCCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCTTTCAACCATCCAGGCTAGAGGAATGCTCCCGGGGAGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCAGTTCAGGCCCTGAGGTTGTCAGGCGAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACAGCCTCTCTGCAACCTTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCCTGAGAAAGAATCAGGGCCCCGTGAGATCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTTCGCGGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCTCTCTGTTTTCAGCAAATCTCCAGCTGCT
GTCCTTGTGCTGATCAGGGCCAACTTCAGGCCCAAACCAACCAAGGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCGAGACAGATGGCCCAAGTGCACAGGCTGGAGGCCAGACAGCTGTCAAGTCTGCTGGACCC
AGGGCCTTTGGATCCTGGAAGAGGACTCTCTGTGAAGCAACTGTGTAAGAAGAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCTGACCCGGCTGATGGCGAGACTCTTTTGGCCCTCAC
CACCAACTACCGTGACAATGTGATCTCCCCGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTTCGG
CAAGGCAGAGGCCACAGAGCTGAGCCCAACAGGCACAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACACGCGCTCCAGCATGCCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCTCAGTTTAGACTTGGCCACCAAGTGCAGCCTCAGGCGTGAAGTGAAGGGGACCCAGGTGG
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGGATGCTGTAACATACTCAGACGCCT
CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAATCTTTGTAACCTCAGTAGCTAG
CGGCGGCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT
TCCTGGCAAACATATGTGGAGCACAAAGGTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGTGGCCCTTCTTGGGTAGCAGGAGTCAGGGGCTGTACCTGGGGGTGCCAGGAATGCTCTCTGACCTAT
CAATAAAGGAAAAAGCAGTAAAAAAGAAAAAAGAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNDNSPAFAESS
LALEIQEDAAPGTLTLLIKLTATDPDQGPNGEVEFFLSKHMPPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNIPIAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNVEEMAGFEFQVIAEDSGQPMLASSVSVWVSLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT
IVARDADSGANGEPLYsirngNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHQKAD IHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLNRNQGNGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGS
RSAIPD TDGPSARAGGQTDPEQEGLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAAAGMKVQGDPGGKTGTEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

Parameter	Value	Unit	Parameter	Value	Unit
α	0.001	cm ² /s	β	0.001	cm ² /s
γ	0.001	cm ² /s	δ	0.001	cm ² /s
ϵ	0.001	cm ² /s	ζ	0.001	cm ² /s
η	0.001	cm ² /s	θ	0.001	cm ² /s
ι	0.001	cm ² /s	κ	0.001	cm ² /s
λ	0.001	cm ² /s	μ	0.001	cm ² /s
ν	0.001	cm ² /s	ξ	0.001	cm ² /s
\omicron	0.001	cm ² /s	π	0.001	cm ² /s
ρ	0.001	cm ² /s	σ	0.001	cm ² /s
τ	0.001	cm ² /s	υ	0.001	cm ² /s
ϕ	0.001	cm ² /s	χ	0.001	cm ² /s
ψ	0.001	cm ² /s	ω	0.001	cm ² /s
Ω	0.001	cm ² /s	Υ	0.001	cm ² /s
Φ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Ξ	0.001	cm ² /s	Θ	0.001	cm ² /s
Λ	0.001	cm ² /s	Σ	0.001	cm ² /s
Π	0.001	cm ² /s	Υ	0.001	cm ² /s
Γ	0.001	cm ² /s	Φ	0.001	cm ² /s
Δ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Θ	0.001	cm ² /s	Ξ	0.001	cm ² /s
Λ	0.001	cm ² /s	Π	0.001	cm ² /s
Σ	0.001	cm ² /s	Υ	0.001	cm ² /s
Φ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Ξ	0.001	cm ² /s	Θ	0.001	cm ² /s
Λ	0.001	cm ² /s	Σ	0.001	cm ² /s
Π	0.001	cm ² /s	Υ	0.001	cm ² /s
Γ	0.001	cm ² /s	Φ	0.001	cm ² /s
Δ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Θ	0.001	cm ² /s	Ξ	0.001	cm ² /s
Λ	0.001	cm ² /s	Π	0.001	cm ² /s
Σ	0.001	cm ² /s	Υ	0.001	cm ² /s
Φ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Ξ	0.001	cm ² /s	Θ	0.001	cm ² /s
Λ	0.001	cm ² /s	Σ	0.001	cm ² /s
Π	0.001	cm ² /s	Υ	0.001	cm ² /s
Γ	0.001	cm ² /s	Φ	0.001	cm ² /s
Δ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Θ	0.001	cm ² /s	Ξ	0.001	cm ² /s
Λ	0.001	cm ² /s	Π	0.001	cm ² /s
Σ	0.001	cm ² /s	Υ	0.001	cm ² /s
Φ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Ξ	0.001	cm ² /s	Θ	0.001	cm ² /s
Λ	0.001	cm ² /s	Σ	0.001	cm ² /s
Π	0.001	cm ² /s	Υ	0.001	cm ² /s
Γ	0.001	cm ² /s	Φ	0.001	cm ² /s
Δ	0.001	cm ² /s	Ψ	0.001	cm ² /s
Θ	0.001	cm ² /s	Ξ	0.001	cm ² /s
Λ	0.001	cm ² /s	Π	0.001	cm ² /s
Σ	0.001	cm ² /s	Υ	0.001	cm ² /s
Φ	0.001	cm ² /s	Ψ		

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGTCTGTCAACTGCCCATACACTTACATGTCTTACTTCTT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTTGGTCCCAACACATT
TCCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGAGAACTGCAGTTCAACTTCTATACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLEFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDAL EELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 174

[illegible]

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTTGGATTGGGGTAGNTT
TTTTTCATGCGCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTCCTC

FIGURE 176

CTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTGAAGTCTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCCGGCTTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAACCA
CAAGGGTGTGTGCTCATCTGCGCAAGGACAGCGCCAGTGCCTGGTGGCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTTCGGGATCAGCCTGGCTGTCTCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGTGGCTGGCCCCCTTTGATGGTGGTGGGAAAGTCTTCACTACCA
TGGGAGCAGCCTGGGGGTTGTGCGCAAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCAGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCTCAATGCCCACCAGCCCAGCACCCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCGAAC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGGTGGGCGAGCCAGTCATTGGCCTGGAGCTGAT
GGTCAACCAACCTGCCATCGGACCCAGCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGT
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAACCTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCAATTGAGCTGCCATGTCCATTGCAGGAATGGGCCATTCC
CCAGCAACTCTTCTCTCTGGTGTGGTGGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCTGG
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACTTGAGCAGCAGGAGCCTGGTGAAGCGGAGGAGCCAGCATGTCTGTTGGCCAGTGTCTCTGTC
TGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGCACGGCCAACCTGTGTGGTGTTCAGCTGCCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCAAGTGAAGTCCCTCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTGATCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCTCCTGGCTGTACTGGCTGGGCTGCTGGTGCTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCCAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCCGCGGGAGGGGCCGGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTAGGTTCC
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGTGGTGTGCGCATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAAC
TCCTCCACCCAACTTCCCCCTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCCATCACCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCCCTTAACCTAGAGGGTCCGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCCCTCTCATCTGACCTTAGTTTGTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTATTAAAAAATATTTGAGAACAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMOKESKENQWL
GVSVRSQGPQGGKIVTCAHRYEARQQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGEWKFC
RPGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGS LADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQ PSTQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYS GVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFFYLILSTSGIS IETTELEVLELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKEYEVTVSNQGQSLRTLGS AFLNIM
WPHEIANGKWL LYPMQVELEGGQGPQKGLCS PRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANC VVFSCLY SFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVI VRANITVKSSIKNLMLRDASTVIPVMVYLDPM AVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYH AVKIPREDRQQFKEEKTGTILRN NWGSPRREGP
DAHPI LAADGHPELGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-
407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408
and 1031-1047

Variable	Mean	SD	Min	Max
Age	38.5	12.5	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	2.5	8	16
Income	1500	500	500	3000
Health status	0.8	0.4	0	1
Exercise frequency	0.3	0.5	0	1
Stress level	0.6	0.5	0	1
Sleep quality	0.7	0.4	0	1
Work satisfaction	0.5	0.5	0	1
Life satisfaction	0.6	0.5	0	1
Depression score	0.4	0.5	0	1
Anxiety score	0.3	0.5	0	1
Quality of life	0.7	0.4	0	1
Healthcare utilization	0.2	0.4	0	1
Health insurance	0.9	0.3	0	1
Chronic disease	0.1	0.3	0	1
Family size	2.5	1.5	1	5
Religious belief	0.5	0.5	0	1
Community involvement	0.3	0.5	0	1
Environmental awareness	0.4	0.5	0	1
Political participation	0.2	0.4	0	1
Cultural engagement	0.3	0.5	0	1
Volunteer work	0.1	0.3	0	1
Charitable donations	0.2	0.4	0	1
Prosocial behavior	0.4	0.5	0	1
Empathy	0.6	0.4	0	1
Altruism	0.5	0.5	0	1
Generosity	0.4	0.5	0	1
Kindness	0.5	0.4	0	1
Helpfulness	0.6	0.3	0	1
Cooperativeness	0.5	0.4	0	1
Compassion	0.6	0.3	0	1
Forgiveness	0.5	0.4	0	1
Patience	0.6	0.3	0	1
Self-control	0.5	0.4	0	1
Emotional stability	0.6	0.3	0	1
Resilience	0.5	0.4	0	1
Optimism	0.6	0.3	0	1
Positivity	0.5	0.4	0	1
Gratitude	0.6	0.3	0	1
Humility	0.5	0.4	0	1
Modesty	0.5	0.4	0	1
Shyness	0.4	0.5	0	1
Introversion	0.3	0.5	0	1
Extroversion	0.4	0.5	0	1
Sociability	0.5	0.4	0	1
Friendliness	0.6	0.3	0	1
Warmth	0.5	0.4	0	1
Openness	0.6	0.3	0	1
Conscientiousness	0.5	0.4	0	1
Neuroticism	0.4	0.5	0	1
Agreeableness	0.5	0.4	0	1
Conformity	0.4	0.5	0	1
Independence	0.3	0.5	0	1
Autonomy	0.4	0.5	0	1
Self-reliance	0.3	0.5	0	1
Confidence	0.5	0.4	0	1
Assertiveness	0.4	0.5	0	1
Leadership	0.3	0.5	0	1
Influence	0.2	0.4	0	1
Power	0.1	0.3	0	1
Control	0.2	0.4	0	1
Authority	0.1	0.3	0	1
Responsibility	0.4	0.5	0	1
Accountability	0.3	0.5	0	1
Integrity	0.5	0.4	0	1
Honesty	0.6	0.3	0	1
Trustworthiness	0.5	0.4	0	1
Reliability	0.4	0.5	0	1
Consistency	0.5	0.4	0	1
Stability	0.6	0.3	0	1
Endurance	0.5	0.4	0	1
Persistence	0.4	0.5	0	1
Perseverance	0.3	0.5	0	1
Fortitude	0.4	0.5	0	1
Bravery	0.3	0.5	0	1
Courage	0.4	0.5	0	1
Valor	0.3	0.5	0	1
Heroism	0.2	0.4	0	1
Altruism	0.3	0.5	0	1
Generosity	0.2	0.4	0	1
Kindness	0.3	0.5	0	1
Helpfulness	0.2	0.4	0	1
Cooperativeness	0.3	0.5	0	1
Compassion</				

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATCCCGAGACGAAGTAGAGGATGATTATTTCCGCACCTGGAGTCCAGGAAAACCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTGACAGTGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTTGAAAG
ATAATTTCTAAGTGAATTTTAAAATAAATAAATTTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAWSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCQPVVYPSVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKTKTLRLPERSRFDTSILPI
CKDSLGMFMNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQODPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
 CTCTTTTCAGCCCGGGATCGCCCCAGCAGGGATGCGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGTCTGCTGCCCTGGGGCGGCCGGCTTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCCTT
 CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCCCTGAAGGCCCTCGCTGGAGATCGAGTACCAAGTTTGA
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAA
 TCAGATGGAGTTTACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTAGCACCATT
 TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAAGCAAAAGTGGGCACATACAAATTTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAACTCCAAACTAGAGTACGTAACTTGAAA
 AATGAGGCATAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAATATTTTGAATATA
 AAAGTAGGAAAACAGGTATAATTTTAAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
 TTGTACTTAAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
 TTTCTTAACCTTTGAAAAATTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
 AGTCTGTTTTTAAACAGGTTCTATTACCCAGAACTTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
 TCAGTTTTTAAGTTATAAAATCACTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCA
 CTTTTCTCTATTACATATGCATCTCTCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
 AGATTTTTTATAACCAATACATTTTCACTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATT
 CCAAAAGCTGACATTTTTCAGATCTTAAATAACACAAAGTTACACTTACTTAAATTAGGACATGTTTTCTCTTTG
 AAATGAAGAATATAGTTTAAAGCTTCTCTCCATAGGGACACATTTTCTCTAACCCCTTAACTAAAGTGTAGGA
 TTTTAAATTAATGTGAGGTAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAAATGATTGCTGACTTGGATAAATTCATTATTACCAGCAGTTATGAAGGAAATA
 TTGCTAAATGATCTGGGCCTACCATAAATAAATATCTCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
 AGAATTTAGAAAAACTTGAGAAAACCTAATCCAAATAAATTAAGTAACTAGTAACTATAAATAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCTTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAATTTGGCTGTAGGTTTTTATTTTCTACAAGAATCTGGTTTGAATTTATTTTGTAAAGCAGGTACATTTTATA
 AAATGTAAGCCTTCTGTAAGTTTACACTGGGTGACATATTTATTAATAAATTTTATTATAACAACCTTTTAT
 TAAATGGCCTTTCTGAACACTTTATTTATTTGATGTTGAAGTAAAGATTAGAAACATAGACTCCCAAGTTTTTAA
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTGAAGTCTATGGGGGTCTTAC
 TCAAGTACTAGTAATTTAAGTCTCATCATGAATGAACATAATTTTAAAGTTATGCCCATTATATAACGTTGTTTTAT
 GACTACATTGTGAGTTAGAAACAACTTAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
 CTTGATGAGCAATAATGATAACAGAGAGTGATTTTCACTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
 TCTTAGGCCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAATAAGAGTAAATGCCGTAT
 ATGATCAATTACCTTAATTTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
 AGGTCAATAAGATCCTTGCCTATGAATACCCCTCCCTTTTGGCGTGTAAATTTGCAATGAGAAGCAAATTTACA
 GTACCATAACTAATAAGCAGGTTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATTTCTA
 CCTCTCCTGTATGGCTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCT
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTTCAGTAACTTCTTACCATAAATAACGATAATTGCTT
 TATTTGGAAGAATTTAGGAATACTAAGGACAATTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
 CATAACCAAAAAAGCAAACTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
 CATATGCTTTTTTTTAAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATGAGTAAGCTGTTTATCACTT
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTT
 CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCCTAGAGTTTGTCTGATATGCATTTGGATGATTAAT
 GTTATGCTGTTCTTTTCACTGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTATGGTAAATTAATCCTTCTTA
 CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATTTGTATGTCTCCTCAAATGAAGATTCTTTAT
 GTGAAATTTTAAAGACATTGATTCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTG
 CTCAAAGTCTTTTATACTTATAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
 AAATTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLSDSFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL
ILDNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDN
IQESNFDVRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 181

[illegible]

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATAACCAGAT
CTCACCAGAGAGTTCGCAGACACTATGCTGCCCTCCCATGGCCCTGCCAGTGTGTCTTGGATG
CTGCTTTCCTGCCTCATTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCAACAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTAGGGCCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCCTCCCAAACCTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTTCCTTTTTCTTCACCTTCATTTCAAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

८

2

•

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLFVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

protein = 9606001

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTC AACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 188

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTACCAATCCCGTGCGCCGCGG
CTGGGCCCGTCCGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTACCCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTTAC
TTAAATCAGAACTTGCATAAGAAAAGAAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTTGG
AGTGTCCAAAACCTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTTGGTATTTATGATGATGATCCTGAAATCATAAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAATCTGGAGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCACTCCATACAACTGCTTTTGTGCTGGTATTGGCTGGCTGATCACTTTTTTGTTCAAA
AGGAGGAGATTGTTTGTACTTACAGACACGACTCAGGCTTAGTGGCATGTTGTTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCAGATTTTGAAGTACTTTTCGGCAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTATTTTTTTCAATTTTGGAAAAATGAAATTCAAATGATCCTGAGCTGAAAAACT
AAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGAAGTCTCTCTGACCCAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAGAATATGAAATTCATCATGGAAA
GAAGATTCTTATATGATATACTTGCCTTTGCCAAAGAAAGTGTGAATCTCATGTTACCACGCTTGGACCTCAAAA
TTTTCTGCTCAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCTTGGTGCTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTTCATGA
GGGACTCTGTAACATGTATAACATTCAAGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ATCCACCACCTTCAACGAAGTAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCCTTGGCAAGTCTTAATGCCAAGTGGAAAGAATGGCCGACATTAAGTGGATGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCCAAATCAAATAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATCCCTGAGAAT
CTGGGGTCTAGGATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGG
GAAAAATCATGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGTCTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAGCTGGAAAAGTAGACTGTCAAGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCTTATCCAAGCTTTAAGTTTTATTCTACGAAAGAGCAAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAATAATTTGGAAGTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTTTGAATAATGTTGAAGATGAAGAAAAGTTTAAAGAAATTCGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAGAAGGGTCTGCAACCTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGTTTACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAAACAACCTTTAAAAAATATTTAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTGGCTGACCTGAAAAGAGGTAAC
TAGTTTTTGGTCACTTGTTCTCTCTAAAAATGCTATCCCTAACCATATATTTATATTTTCGTTTTTAAACACCCAT
GATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATTTATGAGGAGATTCTTCATTGTTTTCTTCTCTCTCA
AAGGTTGAAAAAATGCTTTTAAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTCTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
TAGCAATTAAGTGGCATTTGATAGATATCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTTCATGTATTTTCTGAAATGCTTTTCATAGAAATTTTCCACTGATAGTTGATTTTTTGGGCATCTAATAT
TTACATATTTGCCCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTCATAGTTTTG
TTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAATTAATTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACGTAGTTATTTCCAGTTACTAGTTTACTGTGAGGGGCTGCCTTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATAAATCTAGGAAAGGATCTTCTAGTTTCTGTGTTGTTTAGA
CTCAAAGAATCACAAATTTGTGAGTAACATGTAGTTGTTTAGTTATAATTCAGAGTGTACAGAATGGTAAAAAT
CCAATCAGTCAAAGAGGTCAATGAATTAAGGCTTGCACCTTTTTTCAAAAAAAAAAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKNRAYEVLKDEDLRKKYDKYGEKGLDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDAAVNSGELWVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSPLYVFQP
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHS
AEQILEFIEDLMNPSVVSILTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPTQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTTACCA
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTTCTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGAAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTACATTTTTT
TCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCCTGTTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCCTT
TCATGCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGA
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT
TCATTAGCCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAGTGGACCCTCTATATTTCTCCCTTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESIVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAACRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDDEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGFTKNPSTRLWPVLETDEVVRSLLD
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 192

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCACAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCAGCCCGCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACAGTCCTTCT
CTAGGCCGCACACGGGGCCGCGCTGCCACGCCCCGGGCCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCGAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCCGCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTCGCGTTGCTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSF SRPHTGPPLPTPGPDRDRELTA
DSDVDEFLLDKFLSAGVKQSDLPRKETEQPPAPGSMEE SVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFP TKERAFDDIPNSEL SHLIVDDRHGAIYCYVPKVACTNWK RVMIVLS
GSL LHRGAPYRDPLRIPREHVHNASAH LTFNKFWRRY GKL SRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFEL ENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKV SFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLD EDAQAQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLF GYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCCGCATGGGCCAGCCCGCTGAGGTGGGGCTGCGGCAGTGTTCTTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGACAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCCCGATATCCCTTCCTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDVNNAGHHPPQRPETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

protein = seqdb2

[illegible]

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACTGT
CTGTTTCTTCTTACCATTTCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACCGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKQGGRPGPLAPGPHQVPLDLVSRMKPYARMEYY

ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCL

GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

Protein: seqdb2

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCCAGCAGACTGATTGAGAATGGGAGGCTTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCCCAGCATTTTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACCTATTTTACAGTGACTCTTGGAATT
CCTGCTTGGTGTCTTATGTGTTTTTTCGTATAGCCACCTTGGTTTTTTGGCCTTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTTAATGATGCGTTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTTGTACCAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTCAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAAACGTAGAAGCCTTGCATCCTTTTTCTGTGTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAATTTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSEWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

protein sequence

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTTCAGGGGCTCTCAAGAAACAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCTCTTAATTGGATTATATATGAGAAGAGCTGTTATCTATTTCAGCATGTCCT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCTTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTACAGTGTGAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACCTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGCGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTGGAGATAGGGTCT
CACTTTGTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCCGCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTAGT
AGAGACAGGGTTTCTCCATGTGCGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA
AAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTTGTCTTATTTTTTTTCATACTTTCCCACTGGTGCTATTTTTATTTCGAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTTCAT
TTCAATAAAAAAAACTCTTATCAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

biochem. biophys. Res. Commun.

[illegible]

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLIQTTFLFLSLLFLVQGAHGRGHREDFRFC SQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QESLAQGPPLLATSVTSWWSPQNISLPSAASF TFSFHSPHTAAHNASVDMCELKRD LQLL
SQFLKHPQKASRRPSAAPASQQ LQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLL VDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLT FQHQLQPKNVT LQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVD AVHKHYLSLLSYVGC VVSALACLVTIAAYLCSRVP LPC
RRKPRDYTIKVHMNLLLAVFLLDTS FLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVS YITNLGLFSLVFLFNMA MLATMVVQILRLRPHTQKWSHVL TLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCTTGGCCTGCCTTGTCACCAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGAGGAAANCNTCGGGACTACACCNCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTGGACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGAGCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCCTTCTGATGGGCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCTGGTGTCTCTGTTCAACATGG

[illegible]

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCTAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACCTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTGCAGCCTTCCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTGCCTGCTTCCAAAGTGCCCGCGCTGAGTGAGCTCTCACCCAGTCAAGCAAATGAGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGCGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCTATGCCACAATTACAGAAAGCTGTGAGTCTTCACTGCTACCCCTTCACTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGAAAGACCTTATTGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACCTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACCTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCTTTCAGTTGAGACCAAGACCCGGTGTGAGGGGATTGCAAAATCACTACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGTCT
TCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCATAAGGACAGGAGAAAAGGTTCTCAATCGTGGAAGAAAATTAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCTGGGCCATAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGTCTCATATTCACAT
ATGTAACCCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAACCTTGT
GTCATGCTGATAGGACAGACTGGATTTTTTCATATTTCTTATTAATAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATTTCTCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTAAATATATCT
ATTTTTTACCAAAGGTATTTAATATTTCTTTTATGACAACTTAGATCAACTATTTTTTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAACTGCAATTTTAAAAAAGTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAATTTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTGAGATCCAGCCATTACTAACCTATTCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCTGATAAAGCGTGTGTGCTGTGCTGTGAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAAATACATGGATATTTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAAGAAAATCAGTAAAATATTTTGTCTGT
AAAATGCTTAATATNGTGCCTAGGTATGTGGTGAATTTTGAATCAAAAATGTATTGAATCATCAATAAAAAGA
ATGTGGCTATTTTGGGAGAAAATTAATAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 206

FIGURE 207

MSLFGLLLLT SALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKSRVVDLNNLLTEEVRLY
SCTPRNFVSISIREELKRDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

protein sequence

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
 AGACTAAAAATGGTGTTCCTAATGTGGACACTGAAGAGACAAATCTTATCCTTTTAAACATAATCCTAATTTCC
 AAACCTCCTTGGGGCTAGATGGTTTCCCTAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
 ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTTCCACGAACACCACGAACCTCACCCCTC
 ACCATTAACCACATACCAGACATCTCCCAGCGTCCTTTACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
 TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
 TTTAGTGGACTCACTTATTTAAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCCG
 CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
 AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTATGTTTCATATTCAATAGAGAAA
 GATGCCTTCTTAAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCCTACTGTT
 TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAAATAAC
 CTCAACCAATTACAAATTCCTTGACCTAAGTGGAATTGCCCTCGTTGTTATAATGCCCCATTTCTTTGTGCGCCG
 TGTAAAAATAATTCTCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGTTTACGTCTA
 CACAGTAACCTCTTTCAGCATGTGCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACGGATCTGTCC
 CAAAACCTTCTTGGCCAAAGAAATTTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
 TCTTTCAATTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
 AAAATTCGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
 AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAAATGCTTAACCTCAGCATGTTTAAACAATTTAAAGA
 CTGAAAGTCATAGATCTTTTCACTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAAT
 GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCCTGGAACAATTACATTTTTCAGATATGATAAGTATGCA
 AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTTATGTCGTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
 TTGGATCTAAGTAAAAATAGTATATTTTGTCAAGTCTCTGATTTTTCAGCATCTTTCTTCTCAAATGCCTG
 AATCTGTGAGGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTTCAACCTTTAGCAGAGCTGAGATATTTG
 GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCAAACTGGAAGTTCTGGAT
 ATAAGCAGTAATAGCCATTATTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
 CTGCAGAAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
 ACTCTGGAATTCAAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAG
 AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAATTCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT
 ATGCCTCCAAATCTAAGAATCTCTCTTTGGCCAAAATGGGCTCAAATCTTTCAAGTTGGAAGAACTCCAGTGT
 CTAAAGAACCTGGAACCTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC
 AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
 CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCTC
 AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGTCTGGTGG
 GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
 CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
 TCTGTATCTCTCTTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
 TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
 GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA
 CATTTTAAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA
 CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTTAAAGATAGCATTTTAC
 TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTTCAGAAG
 TCCAAGTTCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTCTTGTAGTGGCCAACAACCCGCAAGCTCAC
 CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
 ACGGCTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCCTGGC

e.

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYR
NPCYVSYSEIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDTGNTFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLNFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETDLDSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWVWNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAAACAGAAACATGGAACCATGTTCTCTTC
AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
GAAGCTATCCTTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGTCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC
CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCTCAACCTAAAAAACCTAAGGGAGTTACTGC
TTGAAGACAACAGTTACCCCAAATACCCTCTGGTTTGGCAGAGTCTTTGACAGAACTTAGTCTAATTTCAAAACA
ATATATACAACATAACTAAAGAGGGCATTTCAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAAGTCT
ATTTTAACAAAGTTTGGGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGCAGAAATTTGGAGTTGCTAT
CACTATCTTTCAATTTCTTTTACACGTGCCACCCAACTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAACA
CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAAATTTAACATTACTAGATTTAAGCGGGAAT
GTCCGAGGTGCTTCAATGCCCATTTCCATGCGTGCCTTGTGATGGTGGTCTTCAATTAATATAGATCGTTTTG
CTTTTCAAAACTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAACGATGCTGCCCGCTTAGAAATACTTGACTTGTCTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCATTTAAGAGGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCTGATGCAGCTTCCAAACTTATCGACTATCAACTTGGGTATTAAATTTTATTA
AGCAAATCGATTTCAAACCTTTTCCAAATTTCTCCAATCTGGAATTTATTTACTTGTGAGAAAAAGAAATATCAC
CGTTGGTAAAAGATACCCGGCAGAGTTATGCAATAGTTCTCTTTTCAACGTATATCCGGAAACGACGCTCAA
CAGATTTTGAATTTGACCCACATTCGAACCTTTTATCATTTTCAACCGTCTCTTTAATAAAGCCACAATGTGCTGCTT
ATGGAAGAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTGAAAAATCTTCTGACA
TTGCCTGTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTAAAGTGAAGTGAATTTTCAAGCATTCTCATG
TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGAGTGAAGTGAATTTTCAAGCATTCTCATG
AAGTTCTAGATCTCAGCTATAATTCACACTATTTTCAAGATAGCAGGCGTAACACATCATCTAGAATTTATTCAA
ATTTCAAAATCTAAAAGTTTAAACTTGAGCCACAACAATTTTATACTTTAAGAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTAGAATTAGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAACTGACACGTCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATG
AAGCATTCCTTAATTTGGCAGCGAGTCTCACTGAACATATAAATGATAATATGTTAAAGTTTTTAACTGGA
CATTACTCCAGCAGTTTCTCTGCTCGAGTTGCTTGACTTACGTGGAACAACTACTCTTTTAACTGATAGCC
TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCACCTACCCTCTGGCTTTC
TTTCTGAAGTCAGTAGTCTGAAGCACTTCGATTTAAGTTTCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
AACTAAGACCACCACCAATTTATCTATGTTGGAATACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAAATCCCAGACTGGTAGATGTCTTTGTGCCAGTCCCTG
GGGATCAAAGAGGGGAAGAGTATTTGTGAGTCTGGAGCTAACAACCTTGTGTTTCAAGATGTCACTGCACTGATATTAT
TTTTCTTCACGTTCTTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTGTTTTTACTGGGATGTTT
GGTTTATATATAATGTGTTTATAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACCTTCTATGATG
CTTACATTTCTTATGACACCAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG
AGAGCCGAGACAAAAACGTTCTCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
TCATGCAGAGCATCAACCAAGCAAGAAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACTTTA
AAACAGCTTTTACTTGGCTTTGCAGAGGCTAATGGATGAGAACATGGATGTGATTATTTATCCTGCTGGAGC
CAGTGTACAGCATTTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA
ACCCGAAGGCAGAAAGCTTGTGTTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAATGATTACCGGTATAACA
ATATGTATGTGATTCATTAAGCAATACTAACTGACGTTAAGTCATGATTTGCGCCATAATAAAGATGCAAAAG
GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAAACTTAGTGGTTTAAAAACAACA
TTGCTGGCCACAGTTTGTGAGGGTCAAGAGTCCAGGCCAGCATAACTGGGTCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCACTGGGGTCACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCCAAGGCAGCTTGCTTC
ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAATCAAAAAAGTGAT
ATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACCACAGGTCCACCAGCTCCATGGGAGTGACCACC
TCAGTCCAGGGAAAAACAGCTGAAGACCAAGATGGTGAAGTCTGATTGCTTCACTTGGTTCATCAACTATTTTCCCT
TGACTGCTGTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTCAATATCTAAGAACTTTTGGCACTGTGACTAATGGTCTTAATA
TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTTATGCTGTGGTTGCGTTTCTGAT
TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTATAGTCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTAAAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA
TTGTTAATTGCCATTGCTGTAAATCTTAAATGAATGAATAAAAAATGTTTCATTTTACAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQNDSVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFNLN
KNLRELLLEDNQLPQIPSGLPESLTSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFQNLTLQLRYNLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNLYLVEIVSGAFLTMLPRLEILDLSFNYIKGSPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSERISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDEV
LDLSYNSHYFRIAGVTHHLEFIQNFNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRDLNRLKHIPNEAFLNLPASLTTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLNLS
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAID
NLMQSINQSKKTVFVLTKKYAKSWNFKTA FYLALQRLMDENMDVIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGFLWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGTCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTTCCTG
GAGGAGCAGCTGGGGTCTGTCTCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCCCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGA
AAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTATTATGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

protein sequence

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCACGGGGGCCCGGGAGGGGAAC TGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCGGAGGAGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTCT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGA CTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCTCCTCCCCTTCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCTCAGCTGAGGGAAGGTACGAGCTCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSDGTLCPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 215

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGGAGGGGACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCCTGGAGGAGCAGCT
GGGGTCTGTCTCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCTCCTCCCCTTCTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSadGTLcVpKGgPPrVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKlQLVLAPLHSLASQALEHGLPDpGSLlVHsfQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

Parameter	Value	Unit
Temperature	25.0	°C
Pressure	1.0	atm
Flow rate	1.0	L/min
Concentration	0.1	mol/L
pH	7.0	
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Column	C18	
Mobile phase	Water/Acetonitrile	
Gradient	0-100% ACN in 10 min	
Flow rate	1.0	mL/min
Temperature	30	°C
Wavelength	254	nm
Scan rate	10	nm/min
Integration time	10	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume</		

FIGURE 218

GGTTGCCACAGCTGGTTTATGGGCCCCGACCACTGGGGCCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCTGGGACGGTCCGTTTCTCGCCGTGAGCTGCCGGCCG
AGTTGGGTCTCCGTGTTTTCAGGCCGGCTCCCCCTTCTGCTGCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTCTTCTCTCGTGTAAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGGTGACACGGAAATGGGAGAAACTCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCGCGTCATGATGGCCCCGGCAAAAGGGCATTCTTCTACCTGACCTTTTCTCATCCTGGGGACATGTACAC
TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTTCTGCTGCTCTCTGCCATCCCTGTATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCTGGAGTGATTCTCGGGCGCTACCAGATG
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
TCAAGAATTTCCAGATAAACAACCAGATTGTGAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGCCTCCCATTTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCTGGGTGGGGAAATTGTG
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGCTTCGCCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAATTTGGCTTCTTGAGACATTGAAAGAACTCTGGAAGTCTG
TCTAGAAAGTCTCATTTGCTTCTTACACTCTGGTCCGTCTGCGGACTGACTGGATTTCATACCTTCTCGTGG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGTGTGGCCCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGT
TTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCCAGAGAACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTGTTGTGTT
TAATTAGGGCTATGAGAGATTTTCAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTCTTTCTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTGAGTCACTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCAAGGGTCCCCCTTGGACCCCCCTCTCTCCCTCCAGATCCAGCCCTCCTGCTTGGGGTCTAC
TGGTCTCATTCTGGGGCTAAAAGTTTTTGGAGTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCAAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTACTCCTGCTCCCATAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
ACTATCCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACAGTTCTTCCATTAAAGCCT
CGGCTGAGTGAGGGAAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCT
ATGGTAACCACTGGGGCTTCTTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCAGAGCCACTT
CACCTTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTTCAGGGAAGAAGATTATGT
ATTATATGTGGCTATATTTCTTAGAGCACTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
GCGGTGGGGGAGTGTAACCGGAATTTTCATCTATTTGAAGGCGATTAACTGTGTCTAATGCA

the = 5000

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIIYVFAFNIVYVALKSLKIGFLETLETPGTVLEVLI CFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPEPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 219

FIGURE 220

AAAACCCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

[illegible]

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGCAACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCCGTGAACCTATCCACCATACA
TTTCAGAAAGCCAAGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCCTGGCTCCAAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGCACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCCACAGAGTGCCCCCAGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

Sequence: 58021001

FIGURE 222

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMD NVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPVVR RVKVTVNYPPISEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLI FNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 222

FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

Sequence of

[illegible]

ATGGCTGGTGAACGGCGGGGCGGGGACGGGGACCGGGGCGCGGGCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCGACTCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGA CAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCAGTACCACAGAGACCCATCCCCACAGCTAGCTGCCTTTACA
GAGGCGCTGCATTGCAGTGGCTGGAAAAATCTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGAGCTTTTAC
CAGTTCTCCTGTGGGGGTGATTTCGGAGGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGA AAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCAGCCACTGAGA
GACCTCATTGAGAAGATTGTTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTATGAGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTACCCTCTACATCAGTGCCGACTCTAAGAGGTTCCAAACAGC
AATGTTATCCAGGCTGACACAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAACACAGACTGCCAATGAG
AAAGTGCTCACTGCCCTATCTGGAATACATGGAGGAACCTGGGATGCTGCTGGGTGGCGGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGAGAGCTGCAGGCTCTGGCGCCCTCCATGGA CTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGT CAGAGCTCATCAACCGCACGGAACCAAGCATCTTGAA CAATTAACCTGATCTGGAACCTGGTGCAAAG
ACAACCTCAAGCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCTCTATGGCCTAAGAAG
TCTGTGTGCCGAGGTGGCAGCTTGCACTCTCAACACGATGACGCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTTCCAGACTTTATCTTGAGGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTCCAAACATCTGAATTTGTACAACTTCTTGCCAAAGTTATGGCTGACCAGCTCCGCAAC
CTCCCAGCCGAGACCTGAGCAGTACGCCCCGACAGCTGAATGCTTACTTCTTCCAAC TAAGAATGAGATC
GTCTTCCCCGCTGGCATCTG CAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATC
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAAACACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAAACATTACTGACAACGGGGGGCTGAAG
GCTGCCCTACAATGCTTACAAAGCATGGCTGAGAAGCATGGGAGGAGCAGCAACTGCCAGCGCTGGGCTCACC
AACCACCAGCTCTTCTTCTGTTGGGATTGTGCCAGGTGTGTTCTCGGTCGCGCACACCGAGAGCTCTCAGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAAC TCCCGTGACTTCTGCGG
CACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGCTAGTCCCTCCCCCACCAGGTGCACATGAGTACAGACCCCTCTCAATCACCACATTG
TGCTCTGCTTTGGGGGTGCCCTGCCCTCCAGCAGAGCCCAACCATCTACTGTGACATCTTCCGTGTACCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRD LIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDY YLNRTANEKVL TAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQ LANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEK LLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTA FEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLP TKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNP GQLCEVW

Type II Transmembrane domain:

amino acids 32-57

protein sequence

[illegible]

Downloaded from ascelibrary.org by University of California, San Diego on 06/06/14. Copyright ASCE. For personal use only; all rights reserved.

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCCTTGTTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTCCGA
GGAACCCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCCT
CCTTTCTTTCTTCTTTTCTTCCCTCCCTCCCTTTCTTCCCTTTTCTTCCCTTCTTCCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTCCG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCTCTAG

FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCCGAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCTT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCTATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGGCTACCCCCACCTCACCCCTGTGCTCCTTGTGCCACTGCTCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 228: SEQUENCE

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCCAAAAACACCGAAAAACAAAAAGAGAGAAACAACACCCAACAACCTGGGGTGG
GGGGAAGAAAGAAAGAAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTCGCGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCAGGAAAATGCTGTGTGCTATT
CCCAGATGTGAGGAAAGTAAAAGTTGTGTCAACTTTGCTCCTACTATTTCAGGAAATTAAAT
CTGGCACCGTGACCCCGGACGACGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTGTATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAACTTACAACCTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTTCAGTTCCTGTTTCAATAAGAGTCAATCCACATTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTGTCTCCAACATAATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCCTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCTTTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTCACTAGTCCAAGCCAAAAA
TTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLP SCLPAGQSVDFPWA AVDNMMVRKGDTAVLR CYLED
GASKGAWLN RSSI IFAGGDKWSVDPRVSISTLNKRDYSLQIQNV DVTDDGPYTCSVQTQH TP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFPDVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLSSTTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGTTCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACTTCACTCAGAACCAGGGATGTTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCCGCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCAAAGGGGAACTGGAATAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTGGGGCCTCCTGGAGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGAGTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAGGTGAAAGAGGTGAAAACCTCAGTGTTCC
GTCAGGATTGTTCGGCAGTAGTAACCGAGGCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTTCAAGTGTTCGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGGAAACCTTTTCA
CTTCTCTGCTCCCGAGGTGTCTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLQLSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPVGKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDGLPLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAKGDQGPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGTGLQGGQGRKGESGVPGPA
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 232

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGGCCAAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCG
TGATTCCTTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAACACCATTCCCTGATTTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

104217-3307.001

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGGG
 CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCGCGGTGGTGGTTGGAGGGC
 GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
 TGAATCTCCTTACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCGCGCTGGCTGTG
 CGCTGGGGCGCTGGTGTGGCGGGTGGCTTCTTTCTCCTCGGCTTCTCTTCGGGTGGTTTA
 TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
 GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
 AGGAACAGAACAAAACCTTTCAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCC
 TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTGTCTACCCAAATAAGACTCATCCC
 AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTCTCTCTC
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
 TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTTGTAATTGCCAGATATGGGAAAGT
 TTTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
 CCGACCCTGCTGACTACTTTTGTCTCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCT
 GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
 AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGAATTTGCAGAGGCTGTTGGTCTTCCAA
 GTATTCCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
 TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCTTACAATGTTGGACCTGG
 CTTTACTGGAACTTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
 CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
 CTGGGAGGTACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
 TGTTTCATGAAATTTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
 CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA
 GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
 AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
 TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
 ACTAAAAAAGTCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAATTGGGATCTGG
 AAATGATTTTGAAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTA
 AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
 GAGTTGGTGGAAAAGTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
 AGGAGGGATGGTGTGTTGAGCTAGCCAATTCCATAGTGCTCCCTTTTGATTGTGAGATTATG
 CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
 ATGAAGACATACAGTGTATCATTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
 TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
 TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGAC
 AGGCCTTTTTTATAGGCATGTCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
 ATTCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
 GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
 TTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
 CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAATTTGGTATATTTGAAATAAAGT
 TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQOLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIRYGVFRGNKVKNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANNEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQERGVAAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPFEFGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSLFSVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713